Topics in mathematical finance

Heritage, J. P.

Topics in Mathematical Finance

submitted by

J.P. Heritage

for the degree of Ph.D.

of the

University of Bath

2002

Signature of Author ....

J.P. Heritage

Acknowledgements

I would like to thank my supervisor Chris Rogers for his contribution to this thesis. I have benefitted greatly from his ideas, guidance, and knowledge of mathematical finance.

I would also like to thank my family, in particular my parents, for their support throughout my education.

Summary

In this thesis we consider four problems from the general area of mathematical finance. In Chapter 2 we consider how to price a moving average up-and-out call option in the Black-Scholes model. This option differs from an up-and-out call option in that knock-out occurs when a moving average of the price process reaches the barrier. We study two such options which differ in the precise conditions for knock-out. Approximate prices are obtained for these options in terms of the prices of up-and-out call options.

In Chapter 3 we present an optimal investment/consumption problem. Agents decide how to divide their wealth between investment in an asset which pays stochastic dividends, and consumption. They aim to maximise the utility of their respective consumption streams. The focus of our work is on how a large agent behaves in this environment. The model can also be interpreted in the context of a corporate takeover.

Chapter 4 is concerned with the study of a Markov process. This Markov process arises from a model describing meetings and interactions between a group of agents. We find that the Markov process is close, in some sense, to a dynamical system. Analysis of the dynamical system helps us to understand the behaviour of the Markov process which will typically exist on a large state space.

In the final chapter we model the activity of a firm with an issue of convertible debt. The value of the firm's assets is taken as the fundamental exogenous stochastic process. We find the conditions under which the firm should default, and the conditions under which a bondholder should convert. This behaviour forms a Nash equilibrium. We find that the bondholders do not all convert at the same time. We establish the share and bond prices within the model.

Contents

List of Figures

Chapter 1

Introduction

The remaining chapters of this thesis are each concerned with a problem in mathematical finance. The setting for the first problem, that of Chapter 2, is the Black-Scholes model of a share price. We aim to price a derivative — the moving average barrier option — within this model. Black & Scholes (1973) and Merton (1973) set the foundations of arbitrage pricing theory by finding a formula for the price of a European call option. They show how to hedge the option with a self-financing portfolio and hence determine the price.

Much of option pricing theory is based on two key results (see Harrison & Kreps (1979) and Harrison & Pliska (1981)). The first is that the absence of arbitrage is equivalent to the existence of an equivalent probability measure under which discounted prices are martingales. We shall refer to this equivalent probability measure as a risk-neutral probability. The second is that if markets are complete then this equivalent probability measure is unique. These results enable us to express the price of any option in a complete market as an expectation in the risk-neutral measure. An introduction to arbitrage pricing theory and the Black-Scholes model can be found in Musiela & Rutkowski (1998).

Several authors have studied barrier options since Merton (1973) found the price of a European down-and-out call option. Rubinstein & Reiner (1991) give closed-form expressions for the prices of knock-in and knock-out barrier options, with barriers either above or below the initial share price. Broadie *et al.* (1997) consider barrier options that can only knock-in or knock-out at discretely sampled times, and Carr (1995) looks at options where the barrier feature only becomes active after an initial protection period, and options which depend on whether a second asset reaches a barrier. The particular type of barrier option we study in Chapter 2 knocks out if a moving average of the share price reaches the barrier. The material of this chapter appears in Heritage (2002).

In Chapter 3 we consider a consumption/investment model. Early work of this kind includes Samuelson (1969) and Merton (1969, 1971). In these papers there is a single agent who divides his wealth between investment in assets with stochastic returns, and consumption. The aim is to find the optimal policy for such an agent under various assumptions about the utility function and about the form of the returns on the assets. More recently, Karatzas *et al.* (1987) look at a consumption/investment model under a wide class of asset price processes and utility functions.

In all the above models prices are determined exogenously. Karatzas *et al.* (1990) study a multi-agent consumption/investment model in which prices are determined endogenously; the emphasis of their work is on establishing the existence and uniqueness of an equilibrium in such a setting. We consider a model with endogenously determined prices where there are two groups of agents, perhaps a large homogeneous pool of price-takers, and a small group of 'large' investors who behave differently. The aim of our study is to determine optimal behaviour for the large agents. Our model can also be interpreted as describing a corporate takeover. This chapter appears in Heritage & Rogers (2002).

The material in Chapter 4 is motivated by 'search' models that have appeared in the economics literature (references are given in section 4.1). In these models there is no centralised market in which agents trade, with buyers and sellers automatically matched. Instead pairs of agents meet according to a random matching process. A trade may occur between two agents when they meet. One reason as to why these models have interested economists is that they provide a natural role for fiat money[1]. Without fiat money, two agents can only trade if there is a 'double co-incidence of wants' — each agent must be prepared to buy the good that the other is selling. Barter can then occur. The presence of fiat money can improve the equilibrium (in some sense) as trade can occur when there is only a single co-incidence of wants.

[1] Fiat money is money that has no intrinsic value; its use is as a medium of exchange.

We abstract from the economics and study a Markov process which corresponds to the random matching of pairs agents, with state changes possibly occurring at these meetings.

In the final chapter we consider a firm with an issue of convertible bonds. We take a structural approach in the model, that is the fundamental exogenous process is the value of the firm's assets. The convertible bonds are considered as contingent claims on the firm's assets. The structural approach was used by Merton (1974) in modelling corporate debt. Our aim is to find a Nash equilibrium between the bondholders (who choose when/whether to convert) and the firm (which chooses when/whether to default).

Ingersoll (1977) and Brennan & Schwartz (1977), amongst others, have studied convertible bonds in structural models. One way in which our approach differs is in the treatment of default. We use methods similar to those in Leland (1994) and Leland & Toft (1996) to endogenously find the optimal time for the firm to default. We also allow the bondholders to convert at different times. As a consequence of this we find that the behaviour of the bondholders in the Nash equilibrium differs from that in the papers referenced above.

Chapter 2

Moving Average Barrier Options

2.1 Introduction

The option that we study here is a variant of the up-and-out barrier option, which we shall henceforth refer to as the *standard* up-and-out barrier option.

The moving average barrier option is defined as follows. The time between purchase and expiry $[0, T]$ is divided into intervals $A_1 = [0, \delta]$, $A_2 = [\delta, 2\delta]$, $\ldots$, $A_N = [T - \delta, T]$ of length $\delta = T/N$. The option knocks out if the average of the asset price over any of these intervals exceeds the barrier, otherwise the payoff is equal to that of a call option. Typically N will range from 20 upwards.

The main result of this chapter is to find an approximation for the price of a moving average barrier option in terms of the price of the corresponding standard barrier option. This is done in the next section. The approximation is justified by simulation in section 2.3. In section 2.4 we consider an alternative moving average barrier option where knock-out occurs if the average of the asset price over *any* interval of length δ between purchase and expiry exceeds the barrier.

2.2 Approximating the option price

In the risk-neutral probability, the asset price at time t is given by the process

$$S_t = S_0 \exp\left(\sigma W_t + \left(r - \sigma^2/2\right) t\right).$$

Here r is the riskless interest rate and σ is the volatility of the asset. Throughout W_t will denote a standard Brownian motion. For notational convenience we shall write

$$\mu = r - \frac{1}{2}\sigma^2$$

and

$$X_t = \sigma W_t + \mu t.$$

The moving average option has payoff

$$S_0 \left(e^{X_T} - k\right)^+ 1\left\{\sup_{j \leq N} \frac{1}{\delta} \int_{A_j} e^{X_u} du < e^b\right\}$$

where T is the expiry time, $S_0 k$ is the strike price and $S_0 e^b$ is the barrier. Therefore the time-0 price of the option in our model is

$$E\left[e^{-rT} S_0 \left(e^{X_T} - k\right)^+ 1\left\{\sup_{j \leq N} \frac{1}{\delta} \int_{A_j} e^{X_u} du < e^b\right\}\right]. \tag{2.1}$$

We may assume without loss of generality that $S_0 = 1$ and $T = 1$. The first of these assumptions is trivial; the second is valid as (2.1) can be expressed as a function of the variables $(r/\sigma^2,\ \sigma^2 T,\ k,\ \delta,\ b)$. It is therefore sufficient to consider the expression

$$p = p(\delta, b, k, r, \sigma) = E\left[e^{-r}(e^{X_1} - k)^+ 1\left\{\sup_{j \leq N} \frac{1}{\delta} \int_{A_j} e^{X_u} du < e^b\right\}\right].$$

Using Girsanov's theorem we can make a transformation of the risk-neutral probability measure P to an equivalent one P^0 in which the process X_t becomes driftless. We then have

$$p = E^0\left[M_1 e^{-r}\left(e^{X_1} - k\right)^+ 1\left\{\sup_{j \leq N} \frac{1}{\delta} \int_{A_j} e^{X_u} du < e^b\right\}\right],$$

where the Cameron-Martin martingale M_t is defined by

$$M_t = \exp\left(-\frac{\mu}{\sigma}W_t - \frac{\mu^2}{2\sigma^2}t\right).$$

Let $\Psi(b)$ denote the time-0 price of a standard up-and-out barrier option with barrier e^b. In our model the price of an option is the expected value of the payoff, appropriately discounted. Thus

$$\begin{aligned}\Psi(b) &= E\left[e^{-r}g(X,b)\right] \\ &= E^0\left[M_1 e^{-r}g(X,b)\right],\end{aligned}$$

where $g(X,b)$ is the time-1 payoff of an option with log price path X and barrier e^b. An explicit formula for $\Psi(b)$ is known (see Conze & Viswanathan (1991) for example) and is reproduced in appendix A. We now aim to relate the price of the moving average barrier option to $\Psi(b)$. Firstly we define $\tilde{t}$ to be the time at which the supremum of X_t in the interval $[0,1]$ is attained, and $A_{\tilde{j}}$ to be the interval in which $\tilde{t}$ lies. We make two assumptions which relate $\tilde{t}$ and $A_{\tilde{j}}$:

A1 $A_{\tilde{j}}$ is the interval over which the integral $\frac{1}{\delta}\int_{A_j} e^{X_u}du$ is maximised.

A2 The distribution of $\tilde{t}$ within the interval $A_{\tilde{j}}$ is uniform.

We define the random variable τ by

$$\tau = X_{\tilde{t}} - \log\left(\frac{1}{\delta}\int_{A_{\tilde{j}}} e^{X_u}du\right). \tag{2.2}$$

We can interpret τ as the difference between $X_{\tilde{t}}$ and an average of the process X_t in the interval that $\tilde{t}$ lies. Thus τ depends only on the shape of X_t in the interval $A_{\tilde{j}}$ and not on the values of X_t in this interval. We make the following approximation:

A3 $g(X,b)$ and τ are independent in the risk-neutral probability.

Appendix B shows in more detail why this is a reasonable approximation to make. This enables us to write our approximation $\hat{p}$ to the price of the moving average barrier option in the following way:

$$\begin{aligned}
\hat{p} &= E^0\left[M_1 e^{-r} g(X, b+\tau)\right] \\
&= E^0\left[\Psi(b+\tau)\right] \\
&= \Psi(b) + E^0\left[\tau\right]\Psi'(b) + \frac{1}{2}E^0\left[\tau^2\right]\Psi''(b) + \ldots \qquad (2.3)
\end{aligned}$$

Our objective now is to calculate $E^0[\tau]$. We make one further assumption here.

A4 The process X_t is like a three-dimensional Bessel process either side of the supremum. More precisely, $\sigma^{-1}\left(X_{\bar{t}} - X_{\bar{t}+t}\right)$ and $\sigma^{-1}\left(X_{\bar{t}} - X_{\bar{t}-t}\right)$ are equal in distribution to three-dimensional Bessel processes.

Rogers & Satchell (1991) perform a similar analysis of a Brownian path near the supremum. Now we have enough information to approximate $E^0[\tau]$. We Taylor expand the exponential and then the logarithm.

$$\begin{aligned}
E^0[\tau] &= E^0\left[-\log\left(\frac{1}{\delta}\int_{A_{\bar{j}}} e^{X_u - X_{\bar{t}}} du\right)\right] \qquad (2.4) \\
&= E^0\left[-\log\left(\frac{1}{\delta}\int_{A_{\bar{j}}} 1 + (X_u - X_{\bar{t}}) + \frac{1}{2}(X_u - X_{\bar{t}})^2 + \ldots du\right)\right] \\
&= E^0\left[-\log\left(1 + \frac{1}{\delta}\int_{A_{\bar{j}}} (X_u - X_{\bar{t}})\, du + \frac{1}{2\delta}\int_{A_{\bar{j}}} (X_u - X_{\bar{t}})^2\, du + \ldots\right)\right] \\
&= E^0\Bigg[-\frac{1}{\delta}\int_{A_{\bar{j}}} (X_u - X_{\bar{t}})\, du - \frac{1}{2\delta}\int_{A_{\bar{j}}} (X_u - X_{\bar{t}})^2\, du - \ldots \\
&\quad + \frac{1}{2}\left(\frac{1}{\delta}\int_{A_{\bar{j}}} (X_u - X_{\bar{t}})\, du + \frac{1}{2\delta}\int_{A_{\bar{j}}} (X_u - X_{\bar{t}})^2\, du + \ldots\right)^2 + \ldots\Bigg] \qquad (2.5)
\end{aligned}$$

Assumption A4 enables us to simplify the first integral in (2.5):

$$\begin{aligned}
& E^0\left[-\frac{1}{\delta}\int_{A_{\bar{j}}} (X_u - X_{\bar{t}})\, du\right] \\
&= E^0\left[\frac{1}{\delta}\int_{(\bar{j}-1)\delta}^{\bar{t}} (X_{\bar{t}} - X_u) du + \frac{1}{\delta}\int_{\bar{t}}^{\bar{j}\delta} (X_{\bar{t}} - X_u) du\right] \\
&= E^0\left[\frac{1}{\delta^2}\int_0^{\delta}\left(\int_0^t \sigma R_u du + \int_0^t \sigma R'_u du\right) dt\right]
\end{aligned}$$

where R_u and R'_u are independent standard three-dimensional Bessel processes beginning at zero. Here we have applied assumption A2 that $\tilde{t}$ is uniformly distributed in the interval $A_{\tilde{j}}$ as well as assumption A4. Changing variables in these integrals yields

$$E^0\left[-\frac{1}{\delta}\int_{A_{\tilde{j}}}(X_u - X_{\tilde{t}})\,du\right] = 2E^0\left[\sigma\sqrt{\delta}\int_0^1\left(\int_0^t R_u du\right)dt\right]$$

and so this term is of order $\sqrt{\delta}$. The other terms in (2.5) are at least of order δ and so if we define

$$\eta = 2\int_0^1\left(\int_0^t R_u du\right)dt$$

then we have

$$E^0[\tau] = \sigma\sqrt{\delta}E^0[\eta] + O(\delta).$$

We will discard the order δ terms. An application of Jensen's inequality to (2.4) shows that

$$E^0[\tau] \le \sigma\sqrt{\delta}E^0[\eta].$$

Therefore discarding the order δ terms produces an overestimate. The density function for R_u (see Revuz & Yor (1991) for example) is

$$f_u(y) = 2y^2\exp\left(-\frac{y^2}{2u}\right) / \left(2\pi u^3\right)^{1/2}.$$

By evaluating the relevant integrals we find that

$$\begin{aligned} E[R_u] &= \sqrt{\frac{8u}{\pi}}, \\ E[\eta] &= \frac{32}{15\sqrt{2\pi}}. \end{aligned}$$

Our analysis indicates that $E^0\left[\tau^2\right]$ is of order δ and so we only consider the first two terms in (2.3). Therefore our approximation for the price of the moving average barrier option is

$$\begin{aligned} \hat{p} &= \Psi(b) + \sigma\sqrt{\delta}\frac{32}{15\sqrt{2\pi}}\Psi'(b) \\ &\approx \Psi(b) + 0.851077\sigma\sqrt{\delta}\Psi'(b). \end{aligned} \tag{2.6}$$

2.3 Simulating the option

To test the accuracy of (2.6) I used Monte Carlo simulation with a control variate technique. We outline the method used and then present the details concerning the probability distributions of some of the quantities that are needed.

First the Brownian path X_t was evaluated at the endpoints of the intervals. The probability that a standard barrier option knocks out given these points on the Brownian path can then be calculated. The details of this are given below. Thus we can find the expected payoff of a standard barrier option given the asset price at the endpoints of the intervals. Conditional on the same set of endpoints of the intervals on the Brownian path, for each interval we can find the probability that the average value of the log asset price exceeds the barrier. Again, the details of this are given below. This enables us to approximate the expected payoff of the moving average barrier option given the asset price at the endpoints of the intervals. Taking averages of the expected discounted payoffs over a large number of simulated Brownian paths gives Monte Carlo estimates for the standard barrier option price and the moving average barrier option price. As an exact price for the standard barrier option is known in closed form, the price of the moving average barrier option can be corrected by the error in the standard barrier option price to give the final estimate. For each set of parameter values, the Monte Carlo estimate is based on 10^5 simulations.

If the Brownian path X_t exceeds the barrier at one of the endpoints then we know that the standard barrier option will knock out. Otherwise, for each interval A_j, we calculate the probability that the Brownian path exceeds the barrier inside this interval. The law of X_t in the interval A_j conditioned upon its values at the endpoints of that interval is that of a Brownian bridge. There are several equivalent definitions of a Brownian bridge (see Rogers & Williams, 2000, theorem IV.40.3); we will use the following:

$$X_{(j-1)\delta+t} = X_{(j-1)\delta} + \left(X_{j\delta} - X_{(j-1)\delta}\right)\frac{t}{\delta} + \sigma\left(1-\frac{t}{\delta}\right)B\left(\frac{\delta t}{\delta - t}\right), \qquad t \in [0,\delta]$$

where $B(t)$ denotes a standard Brownian motion. Therefore

$$\begin{aligned}
& P\left(\sup_{0\le t\le\delta} X_{(j-1)\delta+t} > b\right) \\
= \; & P\left(\sup_{0\le t\le\delta} \left(X_{j\delta} - X_{(j-1)\delta}\right)\frac{t}{\delta} + \sigma\left(1-\frac{t}{\delta}\right)B\left(\frac{\delta t}{\delta - t}\right) > b - X_{(j-1)\delta}\right).
\end{aligned}$$

If we make the transformation of variables

$$u = \frac{\delta t}{\delta - t}$$

we find that this probability becomes

$$\begin{aligned}
& P\left(\sup_{0\le u} \left(X_{j\delta} - X_{(j-1)\delta}\right)\frac{u}{u+\delta} + \sigma\frac{\delta}{u+\delta}B(u) > b - X_{(j-1)\delta}\right) \\
= \; & P\left(\sup_{0\le u} -(b - X_{j\delta})u + \sigma\delta B(u) > \delta(b - X_{(j-1)\delta})\right) \\
= \; & \exp\left(-2\frac{(b - X_{(j-1)\delta})(b - X_{j\delta})}{\sigma^2\delta}\right).
\end{aligned}$$

The final equality is a standard result concerning the law of the supremum of a Brownian motion with negative drift.

We also need to find the distribution of the average of the path X_t over an interval given its value at the endpoints. Define Y to be this random variable.

$$Y = \frac{1}{\delta}\int_0^\delta \left(X_{(j-1)\delta} + \left(X_{j\delta} - X_{(j-1)\delta}\right)\frac{t}{\delta} + \sigma\left(1-\frac{t}{\delta}\right)B\left(\frac{\delta t}{\delta - t}\right)\right)dt.$$

As Y is the integral of a linear function of a Brownian motion, it will have the normal distribution. Calculating the mean is straightforward — we take the expectation inside the integral:

$$\begin{aligned}
EY &= \frac{1}{\delta}\int_0^\delta \left(X_{(j-1)\delta} + \left(X_{j\delta} - X_{(j-1)\delta}\right)\frac{t}{\delta} + \sigma\left(1-\frac{t}{\delta}\right)E\left[B\left(\frac{\delta t}{\delta - t}\right)\right]\right)dt \\
&= \frac{1}{2}\left(X_{(j-1)\delta} + X_{j\delta}\right)
\end{aligned}$$

as the expectation inside the integral is zero. To calculate the variance of Y, we first remove the deterministic part and then apply standard properties of variance to deduce that

$$
\begin{aligned}
Var(Y) &= \frac{\sigma^2}{\delta^2} Var\left(\int_0^\delta \left(1-\frac{t}{\delta}\right) B\left(\frac{\delta t}{\delta - t}\right) dt\right) \\
&= \frac{\sigma^2}{\delta^2} E\left(\int_0^\delta \left(1-\frac{t}{\delta}\right) B\left(\frac{\delta t}{\delta - t}\right) dt\right)^2 \\
&= \frac{\sigma^2}{\delta^2} E\left(\int_0^\delta \int_0^\delta \left(1-\frac{t}{\delta}\right) B\left(\frac{\delta t}{\delta - t}\right) \left(1-\frac{s}{\delta}\right) B\left(\frac{\delta s}{\delta - s}\right) dsdt\right).
\end{aligned}
$$

We take the expectation inside the double integral. Using

$$
\begin{aligned}
E\left[B\left(\frac{\delta t}{\delta - t}\right) B\left(\frac{\delta s}{\delta - s}\right)\right] &= \frac{\delta s}{\delta - s} \wedge \frac{\delta t}{\delta - t} \\
&= \begin{cases} \frac{\delta s}{\delta - s} & s \leq t \\ \frac{\delta t}{\delta - t} & t \leq s \end{cases}
\end{aligned}
$$

yields

$$
\begin{aligned}
Var(Y) &= \frac{\sigma^2}{\delta^2} \int_0^\delta \int_0^\delta \left(1-\frac{t}{\delta}\right) \left(1-\frac{s}{\delta}\right) \left(\frac{\delta s}{\delta - s} \wedge \frac{\delta t}{\delta - t}\right) dsdt \\
&= 2\frac{\sigma^2}{\delta^2} \int_0^\delta \int_0^t \left(1-\frac{t}{\delta}\right) s \; dsdt \\
&= \frac{\sigma^2 \delta}{12}.
\end{aligned}
$$

Now that we know the distribution of Y, for each interval we can calculate the probability that the average of the log asset price exceeds the barrier. Hence we can approximate the expected payoff of a moving average barrier option given the asset price at the endpoints of the intervals.

Simulating the average value of the log asset price as opposed to the log of the average value of the asset price introduces an error. Jensen's inequality gives the sign of the error:

$$
\log\left(\frac{1}{\delta}\int_{A_j} e^{X_t} dt\right) > \frac{1}{\delta}\int_{A_j} X_t dt
$$

We can estimate the magnitude of this error using Taylor expansions.

$$
\begin{aligned}
& \log\left(\frac{1}{\delta}\int_{A_j} e^{X_t}dt\right) - \frac{1}{\delta}\int_{A_j} X_t dt \\
= \; & \log\left(\frac{1}{\delta}\int_{A_j} 1 + X_t + \frac{1}{2}X_t^2 + \ldots dt\right) - \frac{1}{\delta}\int_{A_j} X_t dt \\
= \; & \log\left(1 + \frac{1}{\delta}\int_{A_j} X_t dt + \frac{1}{2\delta}\int_{A_j} X_t^2 dt + \ldots\right) - \frac{1}{\delta}\int_{A_j} X_t dt \\
= \; & \left(\frac{1}{2\delta}\int_{A_j} X_t^2 dt + \ldots\right) - \frac{1}{2}\left(\frac{1}{\delta}\int_{A_j} X_t dt + \frac{1}{2\delta}\int_{A_j} X_t^2 dt + \ldots\right)^2 + \ldots \\
= \; & \frac{1}{2}\left\{\frac{1}{\delta}\int_{A_j} X_t^2 dt - \left(\frac{1}{\delta}\int_{A_j} X_t dt\right)^2\right\} + O(\delta^{3/2}) \qquad (2.7)
\end{aligned}
$$

The expected value of the first term in (2.7) is $\sigma^2\delta/12$ which is small for realistic parameter values.

Monte Carlo simulation gives estimates for the prices of the options. We would like to know how accurate these estimates are. We introduce some temporary notation to find the variance of the Monte Carlo estimates: let Z_i denote the discounted payoff for the ith simulated Brownian path, for $i = 1, \ldots, M$, with M denoting the total number of simulations run. The Monte Carlo price, MCP, is given by

$$MCP = \frac{1}{M}\sum_{i=1}^{M} Z_i.$$

As the Z_i are independent and identically distributed,

$$Var(MCP) = \frac{1}{M}Var(Z_1).$$

An unbiased estimator s_Z^2 for $Var(Z_1)$ is

$$s_Z^2 = \frac{1}{M-1}\left[\sum_{i=1}^{M} Z_i^2 - \frac{1}{M}\left(\sum_{i=1}^{M} Z_i\right)^2\right],$$

and so an estimator s^2_{MCP} for the variance of the Monte Carlo estimate is

$$s^2_{MCP} = \frac{1}{M(M-1)}\left[\sum_{i=1}^{M} Z_i^2 - \frac{1}{M}\left(\sum_{i=1}^{M} Z_i\right)^2\right].$$

This is likely to be an overestimate of the variance as it does not take into account the control variate used. Assuming that the Monte Carlo estimate is normally distributed then allows us to find confidence intervals for the Monte Carlo price.

Figures 2-1 to 2-4 show the ratio of the simulated price over the price obtained using (2.6). As the interest rate (figure 2-1) and the strike (figure 2-2) vary, the ratio stays between 0.98 and 1.00 for nearly each point plotted. As the barrier varies (figure 2-3) the ratio is between 0.98 and 1.00 except for low barriers. The range of the ratio is similar when the volatility varies (figure 2-4) except that the ratio is higher when the volatility is high and N is low. A possible reason why the ratio is slightly below one for most choices of parameter values is assumption A1. It is equivalent to assuming that the moving average barrier option knocks out if and only if it knocks out in the interval in which $\tilde{t}$ lies. Thus we underestimate the probability of knock-out and so overestimate the option price, causing the ratio to be below one. With choices of parameter values such as very low barriers or very high volatilities the approximation becomes less reliable. In both these cases the probability that the payoff (for both the standard barrier option and the moving average barrier option) is zero, either because the option knocks out or because the asset value at expiry is less than the strike, is close to one. Therefore the price of both options would be small. The approximation is also less accurate when N is small, corresponding to δ being large. Most of the approximations that we made were based around δ being small and so this effect is to be expected. Finally, the 95% confidence interval for the ratio of any point in figures 2-1 to 2-4 is at most ± 0.01. That is, the 95% confidence interval for the ratio is at most the interval with boundaries 0.01 either side of the point. Observe that these confidence intervals are consistent with the general roughness of the plots.

2.4 An alternative moving average barrier option

Now we consider an alternative moving average barrier option which knocks out if the average price of the asset over *any* interval of length δ exceeds the barrier. The time-1

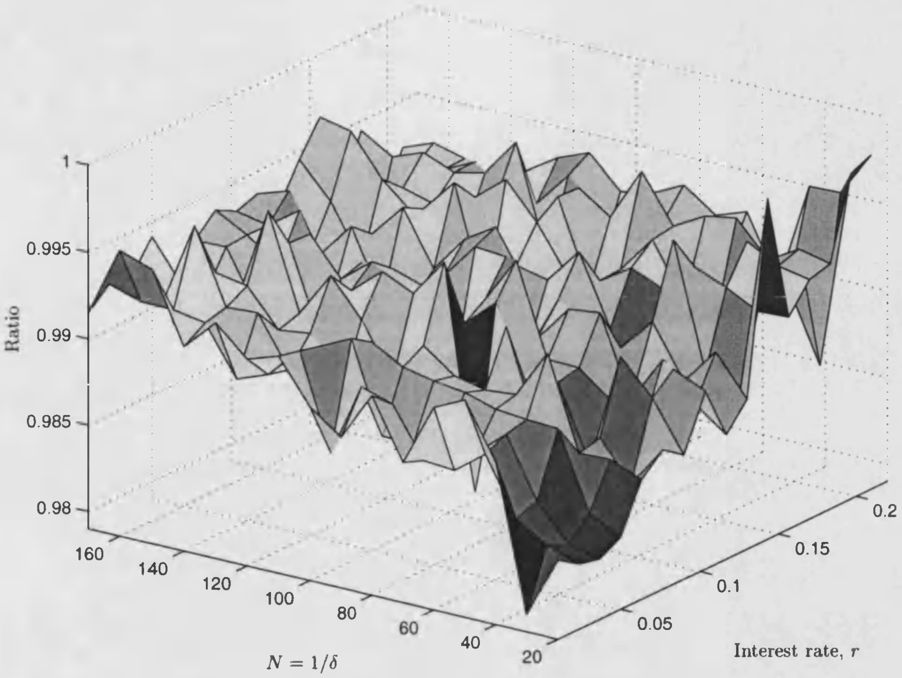

Figure 2-1: Ratio of approximate price to simulated price as the interest rate and δ vary. Other parameter values are $b = 0.3$, $k = 0.9$, $\sigma = 0.25$.

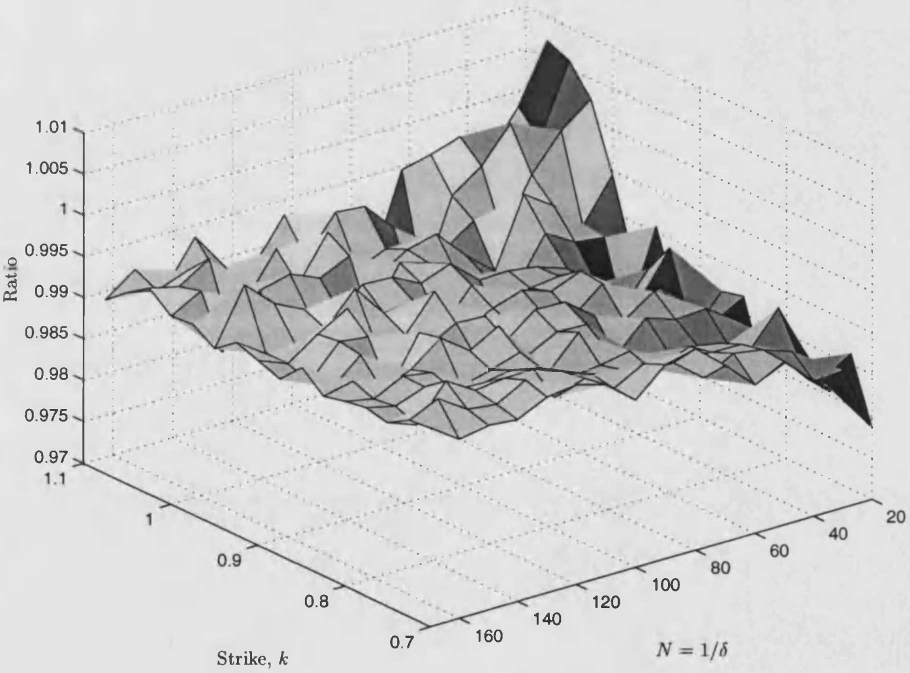

Figure 2-2: Ratio of approximate price to simulated price as the strike and δ vary. Other parameter values are $b = 0.3$, $r = 0.06$, $\sigma = 0.25$.

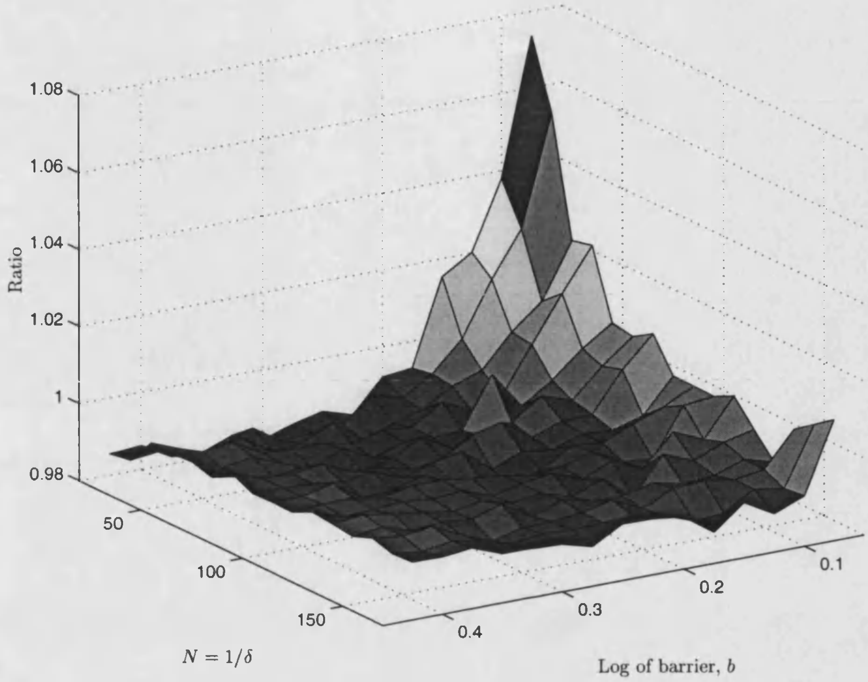

Figure 2-3: Ratio of approximate price to simulated price as the barrier and δ vary. Other parameter values are $\sigma = 0.25$, $k = 0.9$, $r = 0.06$.

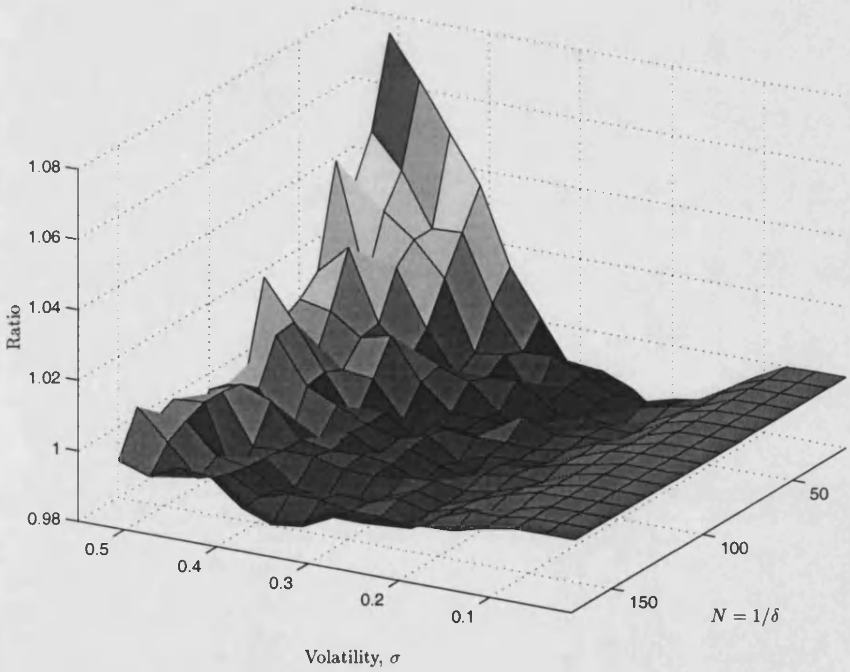

Figure 2-4: Ratio of approximate price to simulated price as the volatility and δ vary. Other parameter values are $b = 0.3$, $k = 0.9$, $r = 0.06$.

payoff of this option is

$$\left(e^{X_1}-k\right)^+ 1\left\{\sup_{\delta\le t\le 1}\frac{1}{\delta}\int_{t-\delta}^{t} e^{X_u}du < e^b\right\}.$$

The payoff, and hence the price of the option, lies between that of the standard barrier option and our first moving average barrier option. We can derive an approximation for the price of this option with a method analogous to that used in section 2.2. The random variable τ would be replaced by

$$\tau' \;=\; X_{\tilde{t}} - \sup_{\delta\le t\le 1}\log\left(\frac{1}{\delta}\int_{t-\delta}^{t} e^{X_u}du\right). \tag{2.8}$$

Let t^* denote the time that the supremum in (2.8) is achieved. A simple differentiation shows that

$$X_{t^*} \;=\; X_{t^*-\delta}$$

and so we may view X_t as a Brownian bridge over the interval $[t^*-\delta, t^*]$. The assumptions made in section 2.2 need to be modified for this option.

A1′ The maximising interval $[t^*-\delta, t^*]$ contains the time $\tilde{t}$.

A2′ $g(X,b)$ and τ' are independent in the risk-neutral probability

We are able to derive an expression for the approximate price $\hat{p}'$ which corresponds to equation (2.3).

$$\hat{p}' \;=\; \Psi(b) + E^0\left[\tau'\right]\Psi'(b) + \frac{1}{2}E^0\left[\tau'^2\right]\Psi''(b) + \dots$$

Now we calculate an approximation for $E^0[\tau']$ in the same way as in section 2.2.

$$\begin{aligned} E^0[\tau'] \;&=\; E^0\left[-\log\left(\frac{1}{\delta}\int_{t^*-\delta}^{t^*} e^{X_u-X_{\tilde{t}}}du\right)\right] \\ &=\; E^0\left[\frac{1}{\delta}\int_{t^*-\delta}^{t^*}\left(X_{\tilde{t}}-X_u\right)du\right] + O(\delta). \end{aligned} \tag{2.9}$$

Vervaat (1979) proved the following relationship between a Brownian excursion and a Brownian bridge.

$$\sigma\sqrt{\delta}Z_{t/\delta} \quad =_d \quad \begin{cases} X_{\tilde{t}} - X_{\tilde{t}+t} & 0 \le t \le t^* - \tilde{t} \\ X_{\tilde{t}} - X_{\tilde{t}-\delta+t} & t^* - \tilde{t} \le t \le \delta \end{cases}$$

Here $=_d$ denotes equality in distribution, and Z_u is a standard scaled Brownian excursion of length 1. A consequence of this result is that the distribution of the integral in (2.9) is the same as the integral of a Brownian excursion.

$$\begin{aligned} E^0[\tau'] &= E^0\left[\frac{1}{\delta}\int_0^\delta \sigma\sqrt{\delta}Z_{u/\delta}du\right] + O(\delta) \\ &= E^0\left[\sigma\sqrt{\delta}\int_0^1 Z_u du\right] + O(\delta). \end{aligned}$$

The density function for Z_t (see Itô & McKean (1996) for example) is

$$f_t(y) \quad = \quad 2y^2 \exp\left(-\frac{y^2}{2t(1-t)}\right) / (2\pi t^3(1-t)^3)^{1/2}.$$

A simple integration then gives

$$\begin{aligned} E[Z_t] &= \sigma\sqrt{\delta}\sqrt{\frac{8t(1-t)}{\pi}}, \\ E^0[\tau'] &= \sigma\sqrt{\delta}\sqrt{\frac{\pi}{8}}. \end{aligned}$$

Therefore the analogous approximation $\hat{p}'$ for the price of this option is

$$\begin{aligned} \hat{p}' &= \Psi(b) + \sigma\sqrt{\delta}\sqrt{\frac{\pi}{8}}\Psi'(b) \\ &\approx \Psi(b) + 0.626657\sigma\sqrt{\delta}\Psi'(b). \end{aligned}$$

As we would expect, the approximation for the price of this option is between that of the standard barrier option and the moving average barrier option considered in section 2.2. Simulation is more difficult in this case as the Brownian path would need to be evaluated much more fully in order to find the supremum of the average price.

2.5 Conclusions

The option that we have studied is difficult to price. The main reason for this is the behaviour of the moving average price process. This process is not Markov; at time t it is a function of price process over the interval $[t - \delta, t]$. It seems that it is not possible to reduce the dimensionality of this dependence. Therefore it is not possible to apply standard partial differential equation methods to pricing the option. Using Monte Carlo simulation to price the option is also not straightforward. The method we used in section 2.3 does give an approximate Monte Carlo price which we expect to be accurate for typical parameter values.

We have used probabilistic methods to find an approximate price for both of the options considered. The results in section 2.3 give an indication as to the accuracy of the approximate price that we found for the first of the moving average barrier options. We would expect that our approximation for the price of the moving average barrier option considered in section 2.4 will depend on the parameter values in a similar way that the first moving average barrier option does.

The accuracy of the assumptions and approximations we have made in section 2.2 depend on the parameter values. This gives us some idea of when we expect the approximate prices we obtain to be close to the true price. Comparing our approximation with the Monte Carlo prices does seem to agree with this, although this effect is complicated by the fact that the accuracy of the Monte Carlo price also depends on the parameters used.

Chapter 3

Large Investors, Takeovers, and the Rule of Law

3.1 Introduction

It is often observed that the assumptions of the Black-Scholes paradigm are all violated in practice, among them the assumption that agents act as price-takers. Prices can be influenced by positions taken, and it is a natural question to ask how this effect operates, say in the simplest situation of two groups of agents, perhaps a large homogenous pool of price-takers, and a small group of 'large' investors who behave differently. Effects of this kind have been studied in examples where there is a group of agents who are following some trading program, as in Frey & Stremme (1997), Gennotte & Leland (1990), Platen & Schweizer (1998), and Brennan & Schwartz (1989). The motivation for the difference in behaviour of this set of agents can be that they are following a portfolio insurance strategy. For example, they seek to trade in the underlying asset in order to dynamically hedge derivative contracts. Thus their behaviour does not fit into the model of a standard utility maximising agent. The paper of Frey & Stremme is typical, in that there are 'reference' traders and 'program' traders, each with their own demand functions, which depend on time, current price, and some 'economic fundamental' process.

This is certainly one approach to the rationalisation of the demand as a function of environment, where the demand functions are in effect given exogenously and prices are derived from that. In this Chapter we take a different approach. We aim to determine the demand by an *endogenous* derivation of optimal investment and consumption paths. There is a single risky asset which we consider as a share in some productive process, generating a dividend stream δ_t. Initially, a pool of $J-1$ agents achieve an equilibrium amongst themselves, while a single (large?) agent J stands aside.

Suppose now that agent J decides to become involved in the market, buying and selling shares in some way that suits his purposes. It is clear that the valuation at time 0 of a share by a member of the pool must depend on the whole of agent J's planned future holdings of shares and not just on the present holding. The share price is the net present value of the future dividend stream from the share and therefore if it were known that J was planning to squeeze the market at some point in the future, the present value of the share would increase. It is clear that we cannot price the share without determining how J will behave in the future.

In the work mentioned above, the behaviour of the program trader is specified exogenously. We attempt to determine the behaviour of J through some optimality criterion. Rather than specifying the number of shares which agent J wishes to hold at times in the future, we shall suppose that J announces at the start that he intends to consume at rate $(1-\varphi_t)\delta_t$ at time t. This leaves a consumption rate of $\varphi_t\delta_t$ for the pool. He chooses φ to maximise his payoff, subject to the constraint that each member of the pool would agree to the proposed deal. To achieve this, he must offer each member of the pool sufficient reward to have an incentive to agree to his proposed φ. We assume that the policy φ is stated, and there may be some initial redistribution of shares, after which the agents in the pool establish an equilibrium based on the declared dividend stream $\varphi\delta \equiv \tilde{\delta}$. We will call the resulting solution the *J-solution* for short.

The situation just described could equally well be interpreted as the takeover of a company (ABC plc), whose shares are held by $J-1$ large shareholders, by another company (XYZ plc) operating an identical technology. XYZ (thought of as J) is assumed to have homogeneous ownership, and makes a proposal to the $J-1$ shareholders of ABC which guarantees them collectively a stated time-dependent deterministic share of the output of the combined firm. The $J-1$ shareholders of ABC then decide between them how this should be divided, according to the equilibrium that they would achieve when faced with the dividend stream $\varphi_t\delta_t$. Alternatively and equivalently, the proposal from XYZ states explicitly what the dividend streams should be for each of the $J-1$ shareholders, in accordance with this equilibrium. The shareholders of ABC now decide whether to accept the offer of XYZ on a take-it-or-leave-it basis.

The theory of this optimal choice is presented in sections 3.2 and 3.3. The model we use is similar to the continuous-consumption model of Bick (1987), where a dividend stream is the fundamental exogenous process and the share price is derived from the optimal behaviour of the agents in an equilibrium. The emphasis in the work of Bick (1987) is to analyse the dynamics of the share price that result in the equilibrium and determine if they are consistent with the Black-Scholes model.

In section 3.4 we investigate a number of examples numerically, comparing the J-solution with the global equilibrium, which would be achieved if J simply entered the market and did not attempt to exploit his power to remain aloof. We find that there is always at least one member of the pool who prefers the global equilibrium to the J-solution; this is a simple consequence of the absence of a blocking coalition for the global equilibrium. Most of the examples are cases where J prefers the J-solution to the global equilibrium. However, an example is given where J prefers the global equilibrium.

The optimal choice for J will of course only work if the rule of law prevails, so that a deal agreed at the start is enforceable. There is nothing in the specification of the J-solution which guarantees that we might not at a later stage find that some subset of the pool might prefer to walk out of the agreed deal and set up a market on their own. Likewise, there is nothing which guarantees that J might not later prefer to take his existing holding of shares and walk out on the deal agreed originally. In section 3.5 we show that in fact the members of the pool will *never* choose to walk out of the original deal, but that J may under certain circumstances prefer to abandon the deal and consume only the dividend from the shares that he currently holds.

3.2 Equilibrium for the pool

3.2.1 The general case

Suppose that the economy consists of a single infinitely-divisible commodity and $J-1$ agents. The supply of the commodity is the dividend of a business which is modelled by the stochastic process δ_t on the time interval $[t_0, \infty)$. This dividend is continuously distributed to each agent at a rate proportional to his share holding, thus at time t agent j receives the commodity at rate $\theta_j(t)\delta_t$, where $\theta_j(t)$ is his share holding. Each agent consumes the commodity on the time interval $[t_0, \infty)$ and aims to maximise the total expected utility of his consumption stream, given by

$$E \int_{t_0}^{\infty} U_j(t, c_j(t))dt,$$

where $c_j(t)$ denotes the rate of agent j's consumption at time t. The utility function $U_j(t, \cdot)$ will be concave and increasing with

$$U_j'(t, 0) \;=\; \infty, \qquad U_j'(t, \infty) \;=\; 0.$$

The total consumption of the agents is determined by a market clearing condition. As agent J consumes at rate $c_J(t) = (1 - \varphi_t)\delta_t$, the consumption rates of the remaining agents must satisfy

$$\sum_{i<J} c_i(t) \;=\; \tilde{\delta}_t \;\equiv\; \varphi_t \delta_t. \tag{3.1}$$

In order to achieve their desired consumption paths, the agents trade the commodity amongst themselves in return for shares or bonds. Both the share price S_t and the bond price B_t are endogenous. The time-t wealth of agent j is defined by

$$w_j(t) \;=\; \theta_j(t)S_t + \xi_j(t)B_t \tag{3.2}$$

where $\xi_j(t)$ denotes the bond holding of agent j at time t. The usual self-financing conditions give the dynamics of the wealth process to be

$$dw_j(t) \;=\; \theta_j(t)\left(dS_t + \delta_t dt\right) + \xi_j(t)dB_t - c_j(t)dt. \tag{3.3}$$

Finally, we require that the wealth process of each agent is always positive. This bounds the consumption of an agent and makes the problem well defined.

We now consider the consumption paths that the agents will follow. The market is complete and so the consumption paths of each agent will satisfy

$$\zeta(t_0, t) \;=\; p_j(t_0)U_j'(t, c_j(t)) \tag{3.4}$$

for some state-price density process $\zeta(t_0, t)$ and positive constants $p_i(t_0), i = 1, \ldots, J-1$, (see, for example, Breeden (1986)). Combining this with market clearing (3.1) gives

$$\tilde{\delta}_t = \sum_{j<J} I_j\left(t, \frac{\zeta(t_0,t)}{p_j(t_0)}\right) \tag{3.5}$$

where $I_j(t,\cdot)$ is the inverse function to $U'_j(t,\cdot)$. The conditions on the utility function mean that (3.5) determines $\zeta(t_0,t)$ uniquely in terms of $\tilde{\delta}_t$ and the constants $p_i(t_0), i = 1,\ldots,J-1$.

The share price is the expected net present value of the dividend stream, thus

$$S_t = \frac{1}{\zeta(t_0,t)} E_t \int_t^\infty \zeta(t_0,u)\delta_u du. \tag{3.6}$$

Similarly the wealth of agent j is the expected net present value of his future consumption stream

$$w_j(t) = \frac{1}{\zeta(t_0,t)} E_t \int_t^\infty \zeta(t_0,u) c_j(u) du. \tag{3.7}$$

Given the processes S_t, $w_j(t)$ and $c_j(t)$, (3.3) determines the share holding process $\theta_j(t)$.

3.2.2 Constant relative risk aversion utility

We now specialise by making the following assumptions, in force for the rest of the Chapter:

(A1) the utility functions are given by

$$U'_j(t,x) = e^{-\rho_j t} x^{-R}$$

for positive constants R and ρ_i, $i = 1,\ldots,J-1$;

(A2) the dividend process is of the form

$$\delta_t = \exp(\sigma W_t + \mu t) \tag{3.8}$$

for some constants σ and μ;

(A3) the process φ is deterministic.

These assumptions are a significant reduction in generality; agents have a common coefficient of relative risk aversion, and differ only in their impatience parameters ρ_i. Stochastic divisions of the dividend process are disallowed. As we shall see, these two assumptions make the problem tractable; allowing different coefficients of relative risk aversion would greatly increase the computational complexity. Although such problems can be handled effectively numerically (see Rogers & Yousaf (2000)), it is not our current purpose to get involved in such complications. The assumption (A2) is probably the least substantive of the three and could be relaxed quite easily, but this seems pointless in view of the other two assumptions being made.

The following proposition summarises the simplifications which result.

Proposition 3.1. *Under assumptions (A1), (A2) and (A3), all agents keep all of their wealth in shares at all times. There are positive constants $p_j(t_0)$, $j = 1, \dots, J-1$ in terms of which the share price may be expressed as*

$$S_t = \delta_t \, \varphi_t^R \, \frac{\psi(t_0,t)}{\tilde{\gamma}(t_0,t)}, \tag{3.9}$$

where

$$\tilde{\gamma}(t_0,t) = \left(\sum_{i<J} \left(p_i(t_0) e^{-\rho_i t}\right)^{1/R} \right)^R$$

and

$$\psi(t_0,t) = \int_t^\infty \tilde{\gamma}(t_0,u) \varphi_u^{-R} e^{\alpha(u-t)} du, \tag{3.10}$$

with

$$\alpha = (1-R)\left(\frac{\sigma^2}{2}(1-R) + \mu\right). \tag{3.11}$$

The optimal consumption streams of the agents are given by

$$c_j(t) = \frac{\left(p_j(t_0) e^{-\rho_j t}\right)^{1/R}}{\tilde{\gamma}(t_0,t)^{1/R}} \tilde{\delta}_t \tag{3.12}$$

and the holdings of shares of agent j at time t is

$$\theta_j(t) = \frac{1}{\psi(t_0,t)} \int_t^\infty \left(p_j(t_0)e^{-\rho_j u}\right)^{1/R} \varphi_u^{1-R} \tilde{\gamma}(t_0,u)^{1-1/R} e^{\alpha(u-t)} du. \quad (3.13)$$

Proof. Assumption (A1) allows the expression for $\zeta(t_0,t)$ in (3.4) to be simplified:

$$\zeta(t_0,t) = p_j(t_0)e^{-\rho_j t} c_j(t)^{-R}. \quad (3.14)$$

Combining this with market clearing (3.1) gives

$$\zeta(t_0,t) = \tilde{\delta}_t^{-R} \tilde{\gamma}(t_0,t). \quad (3.15)$$

Substituting this expression for $\zeta(t_0,t)$ into (3.6) gives the form of the share price S_t in (3.9). The consumption stream of agent j, given by (3.12), is found by eliminating $\zeta(t_0,t)$ from (3.14) and (3.15). The wealth of agent j, given in (3.7), simplifies because of the expressions for $\zeta(t_0,t)$ in (3.15) and $c_j(t)$ in (3.12):

$$\begin{aligned} w_j(t) &= \frac{1}{\zeta(t_0,t)} E_t \int_t^\infty \tilde{\delta}_u^{1-R} \left(p_j(t_0)e^{-\rho_j u} \tilde{\gamma}(t_0,u)^{R-1}\right)^{1/R} du \\ &= \frac{\delta_t \varphi_t^R}{\tilde{\gamma}(t_0,t)} \int_t^\infty \left(p_j(t_0)e^{-\rho_j u} \tilde{\gamma}(t_0,u)^{R-1}\right)^{1/R} \varphi_u^{1-R} e^{\alpha(u-t)} du \\ &= \frac{S_t}{\psi(t_0,t)} \int_t^\infty \left(p_j(t_0)e^{-\rho_j u} \tilde{\gamma}(t_0,u)^{R-1}\right)^{1/R} \varphi_u^{1-R} e^{\alpha(u-t)} du \end{aligned}$$

To find the share holding of agent j, we calculate the dynamics of the wealth process of that agent:

$$\begin{aligned} dw_j(t) &= \frac{dS_t}{\psi(t_0,t)} \int_t^\infty \left(p_j(t_0)e^{-\rho_j u} \tilde{\gamma}(t_0,u)^{R-1}\right)^{1/R} \varphi_u^{1-R} e^{\alpha(u-t)} du \\ &\quad + \frac{S_t \tilde{\gamma}(t_0,t) \varphi_t^{-R}}{\psi(t_0,t)} \left(\int_t^\infty \left(p_j(t_0)e^{-\rho_j u} \tilde{\gamma}(t_0,u)^{R-1}\right)^{1/R} \varphi_u^{1-R} e^{\alpha(u-t)} du \right) dt \\ &\quad - \frac{S_t \varphi_t^{1-R}}{\psi(t_0,t)} \left(p_j(t_0)e^{-\rho_j t} \tilde{\gamma}(t_0,t)^{R-1}\right)^{1/R} dt \\ &= \frac{w_j(t)}{S_t} (dS_t + \delta_t dt) - c_j(t) dt \end{aligned}$$

The expression for $\theta_j(t)$ given in (3.13) follows from this and (3.3). We also deduce that no agent holds any bonds at any time. □

The vector $p(t_0)$ determines the state-price density and the consumption paths, and hence the holdings of shares. In what follows, we have taken as given the initial share holdings $\theta(t_0)$ of the agents, and then computed the values of $p(t_0)$ which match (3.13) to the given $\theta(t_0)$. We shall write

$$\begin{aligned}
\pi_j^{t_0} &= E\int_{t_0}^{\infty} U_j(t, c_j(t))dt \\
&= \frac{1}{1-R}E\int_{t_0}^{\infty} e^{-\rho_j t}\left(\left(\frac{p_j(t_0)e^{-\rho_j t}}{\tilde{\gamma}(t_0,t)}\right)^{1/R} \delta_t\left(1-\theta_J(t_0)\right)\right)^{1-R} dt \\
&= \frac{\left(\left(1-\theta_J(t_0)\right)\delta_{t_0}\right)^{1-R}}{1-R}\int_{t_0}^{\infty} e^{(\alpha-\rho_j)t}\left(\frac{p_j(t_0)e^{-\rho_j t}}{\tilde{\gamma}(t_0,t)}\right)^{(1/R)-1} dt
\end{aligned}$$

for the payoffs of the different agents in the original equilibrium.

3.3 J's optimisation problem

We now consider a Jth agent who will follow the consumption path

$$c_J(t) = (1-\varphi_t)\,\delta_t.$$

The problem for agent J is to choose the function φ_t which maximises his total expected utility of consumption. However, he is constrained in the choice of φ_t. One possible consumption stream is given by taking $\varphi_t = 1-\theta_J(t_0)$. This choice of φ_t does not require any trading with the pool as J is consuming his share of the dividend as he receives it. But for all other choices of φ_t agent J requires the cooperation of the pool in attaining the desired consumption stream. We will suppose that the pool will accept a particular φ_t if each member of the pool prefers, or is indifferent to, that choice of φ_t over taking $\varphi_t = 1-\theta_J(t)$. The preferences of an agent between various proposed functions φ_t are deduced from the relative total expected utility of the corresponding consumption streams. Agent J, therefore, has the following problem

$$\sup_{\varphi,\{p_i(t_0)\}} E\int_{t_0}^{\infty} U_J(t,(1-\varphi_t)\delta_t)dt \tag{3.16}$$

subject to the constraints

$$E\int_{t_0}^{\infty} U_j(t, c_j(t))dt \geq \pi_j^{t_0} \qquad j = 1,\ldots,J-1 \tag{3.17}$$

where $\pi_j^{t_0}$ is the total expected utility agent j obtains when $\varphi_t = 1 - \theta_J(t_0)$.

We can solve this problem with a Lagrangian. The $y_i, i = 1, \ldots, J-1$ are non-negative Lagrange multipliers.

$$\begin{aligned}
L &= E\int_{t_0}^{\infty}\left\{U_J(t,(1-\varphi_t)\delta_t) + \sum_{i<J} y_i U_i(t, c_i(t))\right\}dt - \sum_{i<J} y_i \pi_i^{t_0} \\
\frac{\partial L}{\partial \varphi_t} &= E\int_{t_0}^{\infty}\left\{-\delta_t U_J'(t,(1-\varphi_t)\delta_t) + \sum_{i<J} y_i U_i'(t, c_i(t))\frac{\partial c_i(t)}{\partial \varphi_t}\right\}dt.
\end{aligned}$$

Setting this derivative to zero gives the optimal φ_t. In the case of constant relative risk aversion a solution is given by the roots of

$$e^{-\rho_J t}(1-\varphi_t)^{-R} = \sum_{i<J} y_i e^{-\rho_i t}\frac{\tilde{\gamma}(t_0,t)^{1-1/R}}{(p_i(t_0)e^{-\rho_i t})^{1-1/R}}\varphi_t^{-R}. \tag{3.18}$$

The function φ_t is determined up to the choice of constants y_i and $p_i(t_0)$. These are chosen to give equality in each constraint (3.17) and to maximise (3.16). Typically $\theta_j(t)$ will not be continuous at t_0 and so there will be a reallocation of shares at t_0.

3.4 Numerical results

In this section we present some examples. The Lagrange multipliers can be calculated numerically when the initial conditions of the problem are specified. We will take $t_0 = 0$. In the case of logarithmic utility, the form of δ_t only influences the payoff by an additive constant, and so this is omitted. For non-logarithmic utility, δ_t has to be specified. We choose δ_t to be of the form given by (3.8) and report the value of the constant α, defined in (3.11). The bold typeface indicates the largest payoff for an agent. The final row shows the proportional change to the equilibrium consumption path that would be required to match the J-solution payoff.

Logarithmic utility ($R = 1$)

Agent	1	2	J
ρ	1.2	0.9	1.8
$\theta(0)$	0.35	0.1	0.55
$p(0)$	1	0.1992	
y	0.4086	0.08139	
Equilibrium payoff	**−0.8623**	**−2.439**	−0.3182
J-solution payoff	−0.8728	−2.5312	**−0.3059**
Change	0.98754	0.92061	1.0223

Logarithmic utility ($R = 1$)

Agent	1	2	J
ρ	1.2	0.6	0.38
$\theta(0)$	0.15	0.15	0.7
$p(0)$	1	0.6579	
y	0.4919	0.3236	
Equilibrium payoff	**−1.323**	−3.104	−0.8893
J-solution payoff	−1.529	**−3.086**	**−0.7762**
Change	0.78143	1.0105	1.0439

Logarithmic utility ($R = 1$)

Agent	1	2	3	4	J
ρ	1.0	1.4	0.7	1.9	1.1
$\theta(0)$	0.2	0.1	0.3	0.25	0.15
$p(0)$	1	0.7006	1.045	2.375	
y	1.212	0.8503	1.267	2.879	
Equilibrium payoff	−1.603	−1.618	**−1.606**	−0.6615	−1.719
J-solution payoff	**−1.601**	**−1.617**	−1.610	**−0.6610**	**−1.718**
Change	1.0015	1.0022	0.99697	1.0010	1.0009

$R = 3 \quad \alpha = -0.12$

Agent	1	2	J
ρ	1.5	1.9	2.2
$\theta(0)$	0.35	0.1	0.55
$p(0)$	1	0.02972	
y	0.1755	0.005216	
Equilibrium payoff	**−2.479**	−24.75	−0.7081
J-solution payoff	−2.517	**−24.51**	**−0.7025**
Change	0.99232	1.0048	1.0040

$R = 0.5 \quad \alpha = 0.06125$

Agent	1	2	3	J
ρ	1.1	1.4	1.7	2.0
$\theta(0)$	0.2	0.3	0.3	0.2
$p(0)$	1	1.671	2.047	
y	0.5196	0.8683	1.064	
Equilibrium payoff	**0.8991**	**0.8205**	0.6764	0.4791
J-solution payoff	0.8853	0.8195	**0.6816**	**0.4819**
Change	0.96945	0.99742	1.0152	1.0115

$R = 2.3 \quad \alpha = -0.0962$

Agent	1	2	3	4	J
ρ	0.9	1.1	1.5	1.8	1.4
$\theta(0)$	0.2	0.1	0.15	0.3	0.25
$p(0)$	1	0.2372	0.7980	4.750	
y	0.3277	0.1878	0.1328	2.229	
Equilibrium payoff	**−6.0065**	**−12.682**	−5.6642	−1.9051	−3.1163
J-solution payoff	−6.0154	−12.691	**−5.6618**	**−1.9033**	**−3.1160**
Change	0.99886	0.99949	1.0003	1.0007	1.0001

	$R = 0.8$ $\alpha = 0.07$		
Agent	1	2	J
ρ	0.2	1.4	1.7
$\theta(0)$	0.35	0.49	0.16
$p(0)$	1	22.99	
y	0.9216	2.160	
Equilibrium payoff	**34.77**	3.446	**2.273**
J-solution payoff	33.93	**3.497**	2.263
Change	0.88522	1.0768	0.97826

In each example, the constraints in (3.17) are met with equality. This means that the payoff for an agent in the pool under the J-solution is equal to the payoff which that agent would obtain if J chose φ_t to be given by $\varphi_t = 1 - \theta_J(0)$. If each agent in the pool preferred the J-solution to the global equilibrium, the pool would be a blocking coalition. The absence of blocking coalitions therefore implies that at least one agent in the pool will prefer the global equilibrium to the J-solution.

If each agent in the pool prefers global equilibrium to the payoff obtained when $\varphi_t = 1 - \theta_J(0)$, as in the first example, then the choice of φ_t leading to the global equilibrium satisfies the constraints. Therefore agent J's payoff under the J-solution will be greater than under global equilibrium, as the J-solution gives J his maximum payoff over functions φ_t which satisfy the constraints.

The final example shows that J does not always prefer the J-solution to global equilibrium. Examples of this kind are harder to find. It seems that small changes in the parameter values can result in J prefering the J-solution to the global equilibrium. For this reason characterising the parameters that lead to the J-solution preference for J is difficult.

3.5 Breakdown of the rule of law

In section 3.3 agent J made the choice of φ_t at time t_0 and it was assumed that each agent would follow the consumption path implied by φ_t. In this section we consider whether the deal reached at time t_0 will ever break down at a future time. This would occur if one of two conditions holds, either

1. a subset of the pool prefers to stop trading outside the subset and forms its own equilibrium, or

2. agent J prefers to stop trading and consumes the dividend as he receives it.

The following lemma shows that condition 1 is never satisfied in the case of constant relative risk aversion.

Lemma 3.1. *Without loss of generality assume that $\rho_1 < \rho_2 < \ldots < \rho_{J-1}$. Let each agent have constant relative risk aversion, suppose that the dividend process is of the form given by (3.8), and assume that $\rho_J \neq \rho_1$. Then every subset of the pool contains at least one agent who prefers the J-solution to the subset equilibrium.*

We will prove this lemma in the case where the coefficient of relative risk aversion, R, is not equal to one. The log utility case, where $R = 1$, requires a separate (but analogous) proof because of the different form of the utility function.

Proof. Assume $R \neq 1$. Suppose that A is a subset of the pool in which each agent prefers the subset equilibrium to the J-solution. The consumption path that agent j in this subset follows when A breaks away at time τ is

$$c_j'(t) = \frac{(p_j(\tau)e^{-\rho_j t})^{1/R}}{\tilde{\gamma}_A(\tau,t)^{1/R}}\delta_t\theta_A(\tau-), \qquad t \geq \tau \tag{3.19}$$

where

$$\begin{aligned}\tilde{\gamma}_A(\tau,t) &= \left(\sum_{i\in A}(p_i(\tau)e^{-\rho_i t})^{1/R}\right)^R, \\ \theta_A(\tau-) &= \sum_{i\in A}\theta_i(\tau-).\end{aligned} \tag{3.20}$$

The share holding process which leads to this consumption path is

$$\theta_j(t) = \frac{\theta_A(\tau-)^{1-R}}{\psi_A(\tau,t)}\int_t^\infty \left(p_j(\tau)e^{-\rho_j u}\right)^{1/R}\tilde{\gamma}_A(\tau,u)^{1-1/R}e^{\alpha(u-t)}du \tag{3.21}$$

where

$$\psi_A(\tau,t) = \int_t^\infty \tilde{\gamma}_A(\tau,u)\theta_A(\tau-)^{-R}e^{\alpha(u-t)}du.$$

The consumption path $c_j(t)$ and share holding process that agents in the pool follow under the original choice of φ_t are given by (3.12) and (3.13). For agent j in the subset A to prefer the alternative consumption path (3.19) to that given by (3.12) at time τ we need

$$E_\tau \int_\tau^\infty \frac{e^{-\rho_j u}}{1-R} c'_j(u)^{1-R} du \;\geq\; E_\tau \int_\tau^\infty \frac{e^{-\rho_j u}}{1-R} c_j(u)^{1-R} du$$

which can also be written as

$$\begin{aligned} &\int_\tau^\infty \frac{e^{-\rho_j u/R+\alpha(u-\tau)}}{1-R} \left(\frac{p_j(\tau)^{1/R}}{\tilde{\gamma}_A(\tau,u)^{1/R}} \theta_A(\tau-) \right)^{1-R} du \\ \geq\; &\int_\tau^\infty \frac{e^{-\rho_j u/R+\alpha(u-\tau)}}{1-R} \left(\frac{p_j(t_0)^{1/R}}{\tilde{\gamma}(t_0,u)^{1/R}} \varphi_u \right)^{1-R} du. \end{aligned}$$

Our aim now is to show that this inequality leads to a contradiction. The vector $p(\tau)$ is chosen so that $\theta(t)$ is continuous at $t=\tau$. This means that (3.13) and (3.21) must be equal when $t=\tau$. Using this, our condition is equivalent to

$$\frac{p_j(t_0)}{(1-R)\psi(t_0,\tau)} \;\geq\; \frac{p_j(\tau)}{(1-R)\psi_A(\tau,\tau)}$$

or again

$$\begin{aligned} &\frac{p_j(t_0)}{1-R}\psi(t_0,\tau)^{-1} \\ \geq\; &\frac{p_j(\tau)}{1-R}\left(\int_\tau^\infty \tilde{\gamma}_A(\tau,u)\theta_A(\tau)^{-R} e^{\alpha(u-\tau)} du \right)^{-1}. \end{aligned} \tag{3.22}$$

An expression for $\theta_A(\tau)$ can be found from the expressions for $\theta_j(t)$ in (3.13) and $\tilde{\gamma}_A(\tau,t)$ in (3.20):

$$\theta_A(\tau) \;=\; \frac{1}{\psi(t_0,\tau)} \int_\tau^\infty \tilde{\gamma}_A(t_0,u)^{1/R} \tilde{\gamma}(t_0,u)^{1-1/R} \varphi_u^{1-R} e^{\alpha(u-\tau)} du.$$

We can take $\theta_A(\tau)$ outside the integral in (3.22) and use the expression above to obtain

$$\begin{aligned} &\frac{p_j(t_0)}{1-R}\int_\tau^\infty \tilde{\gamma}_A(\tau,u) e^{\alpha(u-\tau)} du \\ \geq\; &\frac{p_j(\tau)}{(1-R)\psi(t_0,\tau)^{R-1}} \left(\int_\tau^\infty \tilde{\gamma}_A(t_0,u)^{1/R} \tilde{\gamma}(t_0,u)^{1-1/R} \varphi_u^{1-R} e^{\alpha(u-\tau)} du \right)^R. \end{aligned}$$

As this inequality holds for all $j \in A$, it follows that we must have

$$\begin{aligned}
&\frac{\left(\sum_{j\in A}(p_j(t_0)e^{-\rho_j s})^{1/R}\right)^R e^{\alpha(s-\tau)}}{1-R}\int_\tau^\infty \tilde{\gamma}_A(\tau,u)e^{\alpha(u-\tau)}du \\
\geq\ & \frac{\left(\sum_{j\in A}(p_j(\tau)e^{-\rho_j s})^{1/R}\right)^R e^{\alpha(s-\tau)}}{(1-R)\psi(t_0,\tau)^{R-1}} \\
&\left(\int_\tau^\infty \tilde{\gamma}_A(t_0,u)^{1/R}\tilde{\gamma}(t_0,u)^{1-1/R}\varphi_u^{1-R}e^{\alpha(u-\tau)}du\right)^R
\end{aligned}$$

for all $s \geq \tau$. Using the definition of $\tilde{\gamma}_A$ in (3.20), and integrating with respect to s on the interval $[\tau,\infty)$, we deduce that

$$\begin{aligned}
&\frac{1}{1-R}\int_\tau^\infty \tilde{\gamma}_A(t_0,s)e^{\alpha(s-\tau)}ds \\
\geq\ & \frac{1}{(1-R)\psi(t_0,\tau)^{R-1}} \\
&\left(\int_\tau^\infty \tilde{\gamma}_A(t_0,u)^{1/R}\tilde{\gamma}(t_0,u)^{1-1/R}\varphi_u^{1-R}e^{\alpha(u-\tau)}du\right)^R. \qquad (3.23)
\end{aligned}$$

From (3.10), $\psi(t_0,\tau)$ can be written as

$$\psi(t_0,\tau) = \int_\tau^\infty \tilde{\gamma}_A(t_0,u)^{1/R}\tilde{\gamma}(t_0,u)^{1-1/R}\varphi_u^{1-R}e^{\alpha(u-\tau)}\left(\frac{\tilde{\gamma}(t_0,u)^{1/R}}{\tilde{\gamma}_A(t_0,u)^{1/R}\varphi_u}\right)du.$$

Substituting this expression for $\psi(t_0,\tau)$ into (3.23) and rearranging gives

$$\begin{aligned}
&\frac{1}{1-R}\frac{\int_\tau^\infty \tilde{\gamma}_A(t_0,u)^{1/R}\tilde{\gamma}(t_0,u)^{1-1/R}\varphi_u^{1-R}e^{\alpha(u-\tau)}\left(\frac{\tilde{\gamma}(t_0,u)^{1/R}}{\tilde{\gamma}_A(t_0,u)^{1/R}\varphi_u}\right)^{1-R}du}{\int_\tau^\infty \tilde{\gamma}_A(t_0,u)^{1/R}\tilde{\gamma}(t_0,u)^{1-1/R}\varphi_u^{1-R}e^{\alpha(u-\tau)}du} \\
\geq\ & \frac{1}{1-R}\left(\frac{\int_\tau^\infty \tilde{\gamma}_A(t_0,u)^{1/R}\tilde{\gamma}(t_0,u)^{1-1/R}\varphi_u^{1-R}e^{\alpha(u-\tau)}\left(\frac{\tilde{\gamma}(t_0,u)^{1/R}}{\tilde{\gamma}_A(t_0,u)^{1/R}\varphi_u}\right)du}{\int_\tau^\infty \tilde{\gamma}_A(t_0,u)^{1/R}\tilde{\gamma}(t_0,u)^{1-1/R}\varphi_u^{1-R}e^{\alpha(u-\tau)}du}\right)^{1-R}.
\end{aligned}$$

Jensen's inequality tells us that the reverse inequality is also true, and so we must in fact have equality. This is only possible when

$$\begin{aligned}
\varphi_t &= k\frac{\tilde{\gamma}(t_0,t)^{1/R}}{\tilde{\gamma}_A(t_0,t)^{1/R}} \\
&= k\frac{\sum_{i<J}p_i(t_0)^{1/R}e^{-\rho_i t/R}}{\sum_{i\in A}p_i(t_0)^{1/R}e^{-\rho_i t/R}} \qquad (3.24)
\end{aligned}$$

for some positive constant k. For φ_t to lie in the range $[0,1]$ it is necessary that $k < 1$. We are able to show that this form of φ_t contradicts that given in (3.18) by looking at the behaviour when t tends to infinity. If agent 1 is not in the subset A then (3.24) implies that φ_t tends to infinity as t increases, which contradicts (3.18) where φ_t is always in the range $[0,1]$. If agent 1 is in subset A then according to (3.18), φ_t tends to k and so

$$\left(\frac{\varphi_t}{1-\varphi_t}\right)^R \longrightarrow \left(\frac{k}{1-k}\right)^R.$$

By rearranging (3.18) we find that

$$\begin{aligned}\left(\frac{\varphi_t}{1-\varphi_t}\right)^R &= \frac{\sum_{i<J} y_i e^{-\rho_i t/R} p_i(t_0)^{(1-R)/R} e^{\rho_J t}}{\left(\sum_{i<J} p_i(t_0)^{1/R} e^{-\rho_i t/R}\right)^{1-R}} \\ &\longrightarrow y_1 e^{(\rho_J-\rho_1)t}\end{aligned}$$

and as $\rho_J \neq \rho_1$ we have a contradiction. □

Now we look at the second condition. The lemma below gives conditions under which J will break away. It is necessary to show that it is possible for the conditions of this lemma to be satisfied. We do this by presenting an example.

Lemma 3.2. *Let each agent have log utility. If $\rho_J < \rho_1 < \ldots < \rho_{J-1}$, $\rho_1 - \rho_J > \rho_2 - \rho_1$ and $p_2(t_0)/p_1(t_0) - y_2/y_1 < 0$ then agent J will eventually prefer to break away from the J-solution and hold onto his shares consuming the dividend as he receives it.*

Proof. At time τ agent J's share holding is given by

$$\theta_J(\tau) = 1 - \frac{\int_\tau^\infty \tilde{\gamma}(t_0,u)du}{\psi(t_0,\tau)}$$

and so the condition for J to break away at time τ is

$$\log\left(1 - \frac{\int_\tau^\infty \tilde{\gamma}(t_0,u)du}{\psi(t_0,\tau)}\right) - \int_\tau^\infty \rho_J e^{-\rho_J(u-\tau)} \log(1-\varphi_u)\, du > 0. \qquad (3.25)$$

We begin by applying Jensen's inequality to the left-hand-side of (3.25).

$$
\begin{aligned}
&\log\left(\frac{\int_\tau^\infty \frac{\tilde\gamma(t_0,u)}{\varphi_u}(1-\varphi_u)\,du}{\int_\tau^\infty \frac{\tilde\gamma(t_0,u)}{\varphi_u}du}\right) - \int_\tau^\infty \rho_J e^{-\rho_J(u-\tau)}\log(1-\varphi_u)\,du \\
> \; & \frac{\int_\tau^\infty \frac{\tilde\gamma(t_0,u)}{\varphi_u}\log(1-\varphi_u)\,du}{\int_\tau^\infty \frac{\tilde\gamma(t_0,u)}{\varphi_u}du} - \int_\tau^\infty \rho_J e^{-\rho_J(u-\tau)}\log(1-\varphi_u)\,du. \qquad (3.26)
\end{aligned}
$$

Each of the two terms in (3.26) is an average of the function $\log(1-\varphi_t)$. In the case of log utility the form of the optimal φ given by (3.18) simplifies to

$$
\varphi_t \;=\; \frac{\sum_{i<J} y_i e^{-\rho_i t}}{e^{-\rho_J t}+\sum_{i<J} y_i e^{-\rho_i t}}. \qquad (3.27)
$$

We will look at the expression in (3.26) for large values of τ. The condition that $\rho_J < \rho_1 < \rho_2 < \ldots < \rho_{J-1}$ and (3.27) imply that $\log(1-\varphi_t)$ is increasing for large t. We have

$$
\begin{aligned}
\frac{\tilde\gamma(t_0,u)}{\varphi_u} &= \frac{\left(\sum_{i<J} p_i(t_0)e^{-\rho_i u}\right)\left(e^{-\rho_J u}+\sum_{k<J} y_k e^{-\rho_k u}\right)}{\sum_{i<J} y_i e^{-\rho_i u}} \\
&= \frac{p_1(t_0)e^{-\rho_1 u}\left(1+\frac{p_2(t_0)}{p_1(t_0)}e^{-(\rho_2-\rho_1)u}+\ldots\right)e^{-\rho_J u}\left(1+\sum_{k<J} y_k e^{-(\rho_k-\rho_J)u}\right)}{y_1 e^{-\rho_1 u}\left(1+\frac{y_2}{y_1}e^{-(\rho_2-\rho_1)u}+\ldots\right)} \\
&= \frac{p_1(t_0)}{y_1}e^{-\rho_J u}\left\{1+\left(\frac{p_2(t_0)}{p_1(t_0)}-\frac{y_2}{y_1}\right)e^{-(\rho_2-\rho_1)u}+\ldots\right\}
\end{aligned}
$$

for large u, using the condition that $\rho_1-\rho_J > \rho_2-\rho_1$ for the final step. After being normalised, $\tilde\gamma(t_0,u)/\varphi_u$ will give a measure that tends to an average of two exponentials, one with rate ρ_J and the other with rate $\rho_J-\rho_1+\rho_2$, as u tends to infinity. A comparison of the average of $\log(1-\varphi_t)$ under this measure with the average of $\log(1-\varphi_t)$ under an exponential measure with rate ρ_J depends on the sign of $p_2(t_0)/p_1(t_0) - y_2/y_1$. As $p_2(t_0)/p_1(t_0) - y_2/y_1$ is negative, the average of an increasing function under the measure generated by $\tilde\gamma(t_0,u)/\varphi_u$ will be greater than its average under an exponential measure of rate ρ_J. Therefore we conclude that in this case J will eventually prefer to break away. □

It remains to show that it is possible for the ρ_i, the $p_i(t_0)$ and the y_i to satisfy the conditions imposed on them for an optimal φ_t. We do this by presenting an example where the Lagrange multipliers have been numerically calculated. The initial time t_0 is taken to equal 0.

Agent	1	2	3	J
ρ	1.2	1.3	1.9	0.1
$\theta(0)$	0.2	0.25	0.2	0.35
$p(0)$	1	1.298	1.316	
y	1.760	2.417	2.287	

In this case J initially prefers to continue with his original choice of φ_t. However, by time 6 he would benefit from holding onto his shares and consuming the dividend as he receives it.

Time	Payoff from original φ_t	Payoff from holding shares
0	**−2.241**	−3.732
3	**−0.1065**	−0.1088
6	−0.003403	**−0.003378**

The function φ_t decreases to zero and $\theta_J(t)$ increases to one. However, agent J is always consuming at a lower rate than he is receiving the dividend. Therefore consuming the dividend as he receives it results in an immediate increase in the consumption rate.

3.6 Conclusions

We have investigated the impact on a simple market of a large investor who does not act as a price taker. Traditional approaches to the effect of a large investor on price have assumed that price is determined by some instantaneous equalising of supply and demand, but, as Arrow & Kurz (1970, p74) have made clear in a somewhat different context, '... we may say that it requires the future to determine the present resource allocation.' It is such an analysis we have conducted here, allowing the large investor to choose a future dividend flow consistent with the current division of the asset among market participants. This can equally be considered to be the problem facing XYZ plc in its attempts to take over ABC plc; the bidder must offer each of the existing shareholders a deal that would leave them no worse off in order to get the offer accepted. This can be compared with the solution that would obtain if the large agent simply entered the market, and allowed a global equilibrium to establish itself. Examples show that the large agent sometimes prefers one, sometimes the other. The reason is that when the large agent sets up an agreed deal with the other agents, he must ensure that they are all no worse off, and even though he may configure the deal optimally for himself subject to this constraint, in a global equilibrium, it may turn out that some of the other agents do worse off than originally, and this may result in the large agent actually preferring the global equilibrium.

Having decided this, we investigated the viability of the large agent's optimal deal in circumstances when there was no enforceability of the deal. It may be that at all times after the deal is set up, all the agents prefer to continue with the deal than to go off in a subset and follow their own equilibrium in that subset. We have only partial results here; we have been able to show that no coalition of the original pool of agents would ever want to walk out on the deal that they agreed to, but that circumstances can arise where the large agent may wish to walk out with his current share holding and consume the output of that. The chief characteristic of the situation where we were able to show this 'walk-out' is that the large agent is very patient. As time increases, his share of the productive asset increases, as does his consumption stream, but it can be that his consumption stream at large time is less than the consumption stream that would accrue from his current holding of shares. This leads him to walk away. The result here is similar to that found in Kydland & Prescott (1977).

The analysis of the unenforceable situation is still far from complete and appears to be difficult; without the rule of law, some of the deals that XYZ would propose would not be agreed by the ABC, because at some later stage XYZ would walk away from the deal. This would change the nature of the optimal solution proposed by XYZ in the first place.

Chapter 4

Mixed Markov Models

4.1 Introduction

It is commonly assumed in financial models that an investor wanting to sell an asset is able to immediately find an investor wanting to buy the asset. For example, the justification of the Black-Scholes option pricing formula in terms of a hedging portfolio depends upon an investor being able to continuously and instantly readjust his portfolio. Similarly, in the consumption/investment model that we considered in Chapter 3 we assumed that the agents could trade with each other in continuous time without any delay occurring between the decision to trade and the trade occurring. In many markets this may be a close approximation to reality. However, there are markets in which the matching of buyer to seller is less easy. A seller and a buyer may have to search for each other and may incur costs during this search.

The above considerations have led several people to study so-called 'search-theoretic' models. Agents may only trade with each other when they meet; typically pairs of agents meet at the jump times of a Poisson process. This type of model has been used to study the effects of the presence of money. Examples of such work include Diamond (1984), Kiyotaki & Wright (1991, 1993), Shi (1995) and Trejos & Wright (1995). Duffie *et al.* (2000) look at issues of price formation and bid-ask spreads in a similar search-theoretic model.

In this chapter we study a simple Markov process which is motivated by these search-theoretic models. Suppose there are a finite number of particles (or agents), each being in one of a finite number of states. For any ordered pair of particles (a, b), particle a 'meets' particle b at constant intensity which does not depend upon the pair selected. After such a meeting, particle a may change state. The probability distribution of the state of particle a after the meeting depends upon the previous state of a, and on the state of b. The Poisson processes that generate the meetings between different ordered pairs are assumed independent. Independence is also assumed between the state changes of the particles in different meetings. This gives rise to a Markov process.

The remainder of this chapter is set out as follows. In section 4.2 we formally define the model. We then go on in section 4.3 to look at the dynamical system which is the limit of our Markov process as we let the number of agents increase.

4.2 The model

Suppose that we have N particles in total. At any time each of these particles is in one of J states. We denote by $n_i(t)$ the number of particles in state i at time t, where $i = 1, \ldots, J$. The process

$$n(t) \quad = \quad (n_1(t), n_2(t), \ldots, n_J(t))$$

describing the number of particles in each state at time t will be our Markov process. The state space of this Markov process is

$$\left\{ a \in \mathbb{N}^J : \sum_j a_j = N \right\}.$$

The allowed jumps of the process are of the form

$$n \quad \rightarrow \quad n - e_i + e_k$$

where $i \neq k$ and e_i denotes a vector with ith component 1, all other components 0. This jump corresponds to a particle changing from state i to state k. We model the rate at which this happens as

$$\frac{1}{N}\sum_j n_i\,(n_j - \delta_{ij})\,q_{ik}^{(j)} \tag{4.1}$$

for some non-negative $q_{ik}^{(j)}$, $i \neq k$, and where δ_{ij} denotes the Kronecker delta. The interpretation of this is that, for each j, the rate at which particles in state i meet those in state j is proportional to $n_i(n_j - \delta_{ij})$, and the probability that such a meeting results in the state i particle jumping to state k is proportional to $q_{ik}^{(j)}$. The factor of $1/N$ in (4.1) can be thought of as scaling time; in section 4.3 we will consider what happens to the sequence of processes we obtain by letting N increase. Summing over all states j gives the total rate.

Note that this definition of the jump rates preserves the following properties which are implied by the interpretation of our model:

- $\sum_{i=1}^{J} n_i(t) = N \quad \forall t,$
- $n_i(t) \geq 0 \quad \forall i.$

It is helpful to define $q_{ii}^{(j)}$ for each i and j by

$$q_{ii}^{(j)} = -\sum_{k\neq i} q_{ik}^{(j)}.$$

Thus the matrices $Q^{(j)}$ defined with (i,k)th element as $q_{ik}^{(j)}$ are Q-matrices.

We make the following assumption for the remainder of the chapter:

$$q_{ik}^{(j)} > 0 \qquad \forall i,j,k \text{ with } i \neq k. \tag{4.2}$$

We will find that it is often more convenient to work with the normalised process $X^N(t)$ defined by

$$\begin{aligned} X^N(t) &= \left(X_1^N(t), X_2^N(t), \ldots, X_J^N(t)\right) \\ &= \left(n_1(t)/N, n_2(t)/N, \ldots, n_J(t)/N\right). \end{aligned}$$

We also define E as

$$E = \left\{ x \in R^N : \sum x_i = 1, x_i \geq 0 \right\}.$$

Now for $f \in C^2(E)$ consider the martingale

$$M_t^{N,f} \equiv f(X^N(t)) - \int_0^t Q^N f(X_s^N) ds$$

where we have

$$\begin{aligned} Q^N f(x) &\equiv \frac{1}{N} \sum_{i,j,k} n_i (n_j - \delta_{ij}) q_{ik}^{(j)} \left[f\left(x + \frac{e_k - e_i}{N} \right) - f(x) \right] \\ &= \sum_{i,j,k} q_{ik}^{(j)} x_i \left(x_j - \frac{\delta_{ij}}{N} \right) N \left[f\left(x + \frac{e_k - e_i}{N} \right) - f(x) \right]. \end{aligned}$$

We Taylor expand f and apply the mean value theorem. Writing ∂_i for $\dfrac{\partial}{\partial x_i}$ we obtain

$$Q^N f(x) = \sum_{i,j,k} q_{ik}^{(j)} x_i \left(x_j - \frac{\delta_{ij}}{N} \right) \left[\partial_k f(x) - \partial_i f(x) + \frac{1}{2N} (\partial_k^2 - 2\partial_k \partial_i + \partial_i^2) f(y(i,k)) \right]$$

for some $y(i,k)$. Since by definition

$$\sum_k q_{ik}^{(j)} = 0,$$

terms in the summand that depend on k only through the factor $q_{ik}^{(j)}$ will vanish. Therefore if we define V by

$$V f(x) = \sum_{i,j,k} q_{ik}^{(j)} x_i x_j \partial_k f(x)$$

we obtain

$$\begin{aligned} Q^N f(x) &= V f(x) - \frac{1}{N} \sum_{i,k} q_{ik}^{(i)} x_i \partial_k f(x) + \frac{1}{2N} \sum_{i,j,k} q_{ik}^{(j)} x_i x_j (\partial_k^2 - 2\partial_k \partial_i + \partial_i^2) f(y(i,k)) \\ &\quad - \frac{1}{2N^2} \sum_{i,k} q_{ik}^{(i)} x_i (\partial_k^2 - 2\partial_k \partial_i + \partial_i^2) f(y(i,k)). \end{aligned}$$

It follows that for a given $f \in C^2(E)$ there exists some constant K_f such that

$$|Q^N f(x) - Vf(x)| \leq \frac{1}{N} K_f. \tag{4.3}$$

Thus the generators Q^N of X^N are 'close' in some sense to V, the generator of the dynamical system X. In the next section we use results of Ethier & Kurtz (1986) to establish the weak convergence

$$X^N \Rightarrow X$$

which the above analysis already suggests.

4.3 Convergence to a dynamical system

We wish to consider the behaviour of this Markov process when N is large. In particular we look at the limit of the Markov process as N tends to infinity. We find that the normalised Markov processes X^N converge in distribution to the dynamical system X. The processes X^N and X all have sample paths in the space $D_E[0, \infty)$. This space contains functions that are right continuous with left limits, with domain $[0, \infty)$ and range E. The dynamical system X is defined by the smooth vector field

$$V(x) = \sum_{i,j,k} x_i x_j q_{ik}^{(j)} e_k \tag{4.4}$$

where e_k denotes the kth unit vector. The next subsection is rather technical and formalises this convergence. We apply results from Ethier & Kurtz (1986).

4.3.1 Proof of convergence

The main result that we wish to apply is corollary 8.6 of Chapter 4 of Ethier & Kurtz (1986). We cannot use this corollary straight away, however; first we need to show that the premises of the corollary are satisfied. Propositions 4.1, 4.2 and 4.3 and Theorems 4.1 and 4.2 help with this.

First note that the state space E is compact. We define $F(t, x)$ to be the integral path of V starting from x at time 0, that is

$$F(t,x) = X(t) \tag{4.5}$$

where

$$\dot{X}(t) = V(X(t))$$

and

$$X(0) = x.$$

The semigroup $(T_t)_{t\geq 0}$ of the process $X(t)$ is given by

$$\begin{aligned}(T_t f)(x) &= E[f(X(t))|X(0)=x] \\ &= f(F(t,x))\end{aligned} \tag{4.6}$$

defined on the Banach space $C(E)$ with the supremum norm. The first line, (4.6), can be taken as the definition of a semigroup for a Markov process.

A semigroup $(T_t)_{t\geq 0}$ on L is *strongly continuous* (see Ethier & Kurtz, 1986, Chapter 1) if

$$\lim_{t\downarrow 0} T_t f = f \quad \text{for every } f \in L$$

and is a *contraction* semigroup if

$$\|T_t\| \leq 1 \quad \text{for all } t \geq 0.$$

We now prove that the semigroup we have defined is in fact a strongly continuous contraction semigroup on $C(E)$. This is necessary in order that we can apply the Hille-Yosida Theorem.

Proposition 4.1. *The semigroup $(T_t)_{t\geq 0}$ defined on $C(E)$ is a strongly continuous contraction semigroup.*

Proof. It is clear that $(T_t)_{t\geq 0}$ is a contraction semigroup: as we are working under the supremum norm we have

$$\|f\| = \sup_{x \in E} |f(x)|$$

and

$$\begin{aligned} \|T_t f\| &= \sup_{x \in E} |f(F(t,x))| \\ &= \sup_{y \in E'} |f(y)| \end{aligned}$$

where E' is some subset of E. Hence we deduce that $\|T_t\| \leq 1$. It remains to prove strong continuity. As f is a continuous mapping of a compact set to itself, f is uniformly continuous. Thus

$$\forall \varepsilon, \exists \delta > 0 \quad \text{s.t.} \quad |x - y| \leq \delta \Rightarrow |f(x) - f(y)| \leq \varepsilon.$$

The vector field V is continuous on a compact set, so it is bounded. Therefore given $\delta > 0$

$$\exists \eta > 0 \quad \text{s.t.} \quad |x - F(x,t)| \leq \delta \quad \forall x, \forall t \leq \eta$$

and the proof is complete. □

We are now ready to apply the Hille-Yosida Theorem (Ethier & Kurtz, 1986, Chapter 1, Theorem 2.6). This theorem, which is reproduced below, tells us that the generator A of the semigroup $(T_t)_{t \geq 0}$, defined by

$$Af = \lim_{t \downarrow 0} \frac{1}{t} (T_t f - f) \tag{4.7}$$

with domain

$$\left\{ f \in C(E) : \lim_{t \downarrow 0} \frac{1}{t} (T_t f - f) \text{ exists} \right\},$$

is dissipative. One more piece of notation: $\mathcal{D}(A)$ denotes the domain of A.

Theorem 4.1 (Hille-Yosida). *A linear operator A on $C(E)$ is the generator of a strongly continuous contraction semigroup on $C(E)$ if and only if:*

- *the domain of A is dense in $C(E)$,*

- *A is dissipative, that is*

$$\|\lambda f - Af\| \;\geq\; \lambda\|f\| \qquad \textit{for all } f \in \mathcal{D}(A) \textit{ and } \lambda > 0,$$

- *the range of $\lambda - A$ is $C(E)$ for some $\lambda > 0$.*

We will now work with A', which is defined to be the restriction of A to $C^2(E)$. Propositions 4.2 and 4.3 and Theorem 4.2 establish the properties that are required to apply corollary 8.6 of chapter 4 of Ethier & Kurtz (1986). Before giving these results we need a definition. A subspace S of the domain of A is said to be a *core* for A if the closure of the restriction of A to S is equal to A. The following proposition (Ethier & Kurtz, 1986, Chapter 1, Proposition 3.3) is used to show that $C^2(E)$ is a core for A.

Proposition 4.2. *Let A be the generator of a strongly continuous contraction semigroup $(T_t)_{t\geq 0}$ on $C(E)$. Let S be a dense subspace of $C(E)$. If $T_t : S \to S$ for all $t \geq 0$ then S is a core for A.*

As $C^2(E)$ is dense in $C(E)$ and, for each $t \geq 0$,

$$T_t : C^2(E) \to C^2(E)$$

we deduce that $C^2(E)$ is a core for A.

We now give proposition 4.3 (Ethier & Kurtz, 1986, Chapter 1, Proposition 3.1). The notation $A|_S$ refers to the restriction of A to S.

Proposition 4.3. *Let A be the generator of a strongly continuous contraction semigroup on $C(E)$. Then a subspace S of $C(E)$ is a core for A if and only if S is dense in $C(E)$ and the range of $\lambda - A|_S$ is dense in $C(E)$ for some $\lambda > 0$.*

We apply this proposition with $S = C^2(E)$. Now we can use the following theorem (Ethier & Kurtz, 1986, Chapter 1, Theorem 2.12):

Theorem 4.2. *A linear operator A' on $C(E)$ is closable and its closure A is the generator of a strongly continuous contraction semigroup on $C(E)$ if and only if:*

- *the domain of A' is dense in $C(E)$*

- *A' is dissipative*
- *the range of $\lambda - A'$ is dense in $C(E)$ for some $\lambda > 0$*

The three criteria of this theorem are satisfied: $C^2(E)$ is dense in $C(E)$; A is dissipative (by Theorem 4.1) and therefore A' is dissipative; and proposition 4.3 proves the final criterion. We deduce that the closure of A' generates a SCCSG on $C(E)$.

We can now verify that the premises of corollary 8.6 of Ethier & Kurtz (1986) chapter 4 hold. The closure of A' generates a SCCSG on the closure of $\mathcal{D}(A')$ (which is $C(E)$). Now the sample paths of X^N and of X are in $D_E[0,\infty)$, and $C(E)$ is clearly an algebra which separates points and is convergence determining. Since E is compact, the compact containment condition (7.9) of Ethier & Kurtz (1986) chapter 3 holds. For any $f \in C^2(E)$, with $g = Vf$ we define

$$\xi_N(t) = f(X^N(t))$$

and

$$\phi_N(t) = Q^N f(X^N(t)).$$

Then (8.8)-(8.10), (8.33) and (8.34) of Ethier & Kurtz (1986) chapter 4 follow immediately. Condition (8.11) of the same chapter is a consequence of (4.3), and so we conclude that

$$X^N \Rightarrow X$$

if

$$X^N(0) \Rightarrow X(0).$$

4.3.2 Fixed points

In this section we show that there is at least one fixed point of the dynamical system. This result does not depend on our assumption in (4.2). First we define the sets F_n for $n \geq 1$ by

$$F_n = \left\{x \in E : F(2^{-n}, x) = x\right\}.$$

(Recall from (4.5) the definition of $F(t, x)$ as the integral path of the vector field starting at x over time t.) Thus for each element x of the set F_n, if the dynamical system is at x at some time t, then it will be at x at time $t + 2^{-n}$. For each $n \geq 1$ the set F_n is compact, and the map

$$x \mapsto F(2^{-n}, x)$$

is a continuous map from E to E. As E is a compact convex set, Brouwer's fixed point theorem (see for example Border, 1985, corollary 6.6) implies that this map has a fixed point. Equivalently, for each $n \geq 1$ the set F_n is non-empty. It follows from the definition of the sets F_n that

$$F_{n+1} \subseteq F_n.$$

Therefore the intersection of any finite collection of the sets F_n is non-empty, which implies that

$$\bigcap_{n \geq 1} F_n \neq \emptyset.$$

Let x^* be an element of this intersection. Then

$$x^* \in F_n \qquad \forall n.$$

We deduce that x^* is a fixed point of the dynamical system by showing that the vector field at x^* is zero.

$$\begin{aligned} V(x^*) &= \lim_{n \to \infty} \frac{F(2^{-n}, x^*) - x^*}{2^{-n}} \\ &= 0. \end{aligned}$$

4.3.3 Fixed points of the system with two states

When there are two states, that is when $J = 2$, it is significantly easier to study this dynamical system. Here we show that there exists a unique fixed point of the dynamical system. We have

$$x_2 = 1 - x_1$$

and

$$\dot{x}_2 = -\dot{x}_1$$

and so the system is in effect one-dimensional. In order to find any fixed points we look at the vector field. It follows from (4.4) that

$$\dot{x}_1 = x_1^2 q_{11}^{(1)} + x_1(1-x_1)\left(q_{11}^{(2)} + q_{21}^{(1)}\right) + (1-x_1)^2 q_{21}^{(2)}.$$

This is a quadratic in x_1 which is easily seen to be positive at $x_1 = 0$ and negative at $x_1 = 1$. There is therefore a unique root in the interval $(0,1)$ corresponding to the unique fixed point of the dynamical system. It also follows that the fixed point is stable.

4.3.4 Fixed points of the system with three states

Having found that there is a unique fixed point when we have two states, we might conjecture that the same is also true when there are three or more states. In this section we look at the case where $J = 3$. As

$$x_3 = 1 - x_1 - x_2 \tag{4.8}$$

we have a two-dimensional process. It also follows that

$$\dot{x}_1 = 0 \text{ and } \dot{x}_2 = 0 \Rightarrow \dot{x}_3 = 0$$

and hence we are looking for roots of $\dot{x}_1 = 0$ and $\dot{x}_2 = 0$. For a root to correspond to a point in E we require

$$(x_1, x_2) \in E'$$

where

$$E' = \{(x_1, x_2) : x_1 \geq 0, x_2 \geq 0, x_1 + x_2 \leq 1\}. \tag{4.9}$$

We apply two different methods to characterise the solutions, that of conic sections and index theory.

Conic sections

Consider the expression for $\dot{x}_1$ obtained from (4.4). If we use (4.8) to eliminate x_3 from this expression then we are left with a quadratic form in x_1 and x_2. Therefore the roots to $\dot{x}_1 = 0$ in the (x_1, x_2) plane will lie on a conic section. Typically this will be a hyperbola, parabola or ellipse. However, if the quadratic form is degenerate then the roots may instead lie on a line, a pair of lines, or be a single point.

Now consider the expression for $\dot{x}_2$ in (4.4). The same argument as above shows that if we eliminate x_3 from this expression, then the roots of $\dot{x}_2 = 0$ are a conic section in the (x_1, x_2) plane. Therefore the roots of the pair of equations $\dot{x}_1 = 0$ and $\dot{x}_2 = 0$ lie on the intersection of two conic sections in the (x_1, x_2) plane. In section 4.3.2 we proved that there is at least one fixed point of the dynamical system. Therefore our two conic sections must have at least one point of intersection in E'. We attempt to find an upper bound on the number of points of intersection.

If the two conic sections are distinct then there can be at most four points of intersection in the (x_1, x_2) plane, and so at most four points of intersection in E'. If the two conic sections are identical, for example they correspond to the same hyperbola, then the intersection will be an infinite set of points. We now aim to show that we can not have an infinite intersection. Lemma 4.1 (which holds for any J, not just $J = 3$) will be used for this purpose.

Lemma 4.1. *Under assumption (4.2) there are no fixed points of the dynamical system on the boundary of E.*

Proof. The points x on the boundary of E are characterised by the fact that at least one component of x is zero. Therefore by symmetry it is sufficient to show that

$$x_1 = 0 \;\Rightarrow\; \dot{x}_1 > 0 \tag{4.10}$$

on E. From (4.4) we obtain

$$\dot{x}_1 = \sum_{i,j} x_i x_j q_{i1}^{(j)}.$$

By the definition of the $q_{ik}^{(j)}$, the only chance for a negative term in this sum occurs when $i = 1$. However, the presence of the factor x_i means that when $i = 1$ the term in the sum will be zero. Our assumption in (4.2) and the fact that $\sum x_i = 1$ on E imply that there must be a strictly positive term in the sum. Therefore (4.10) holds. $\square$

Suppose that our two conic sections are the same hyperbola, parabola, or contain the same line. We know that a fixed point exists in E so our intersection must include a point in E'. By the nature of hyperbolae, parabolae and lines, it would also include a point on the boundary of E', contradicting lemma 4.1. Therefore the only way we could have an infinite intersection would be if both conic sections were the same ellipse lying entirely in the interior of E'. To rule out this possibility we look at the quadratic form corresponding to $\dot{x}_1 = 0$. We do not need to find all the coefficients explicitly; some are abbreviated to $(\ldots)$.

$$\begin{aligned}\dot{x}_1 &= x_1^2\left(q_{11}^{(1)} - q_{11}^{(3)} - q_{31}^{(1)} + q_{31}^{(3)}\right) + x_1x_2(\ldots) + x_2^2(\ldots) \\ &\quad + x_1\left(q_{11}^{(3)} + q_{31}^{(1)} - 2q_{31}^{(3)}\right) + x_2(\ldots) + q_{31}^{(3)}.\end{aligned}$$

When $x_1 = x_2 = 0$ we have $\dot{x}_1 = q_{31}^{(3)} > 0$. When $x_1 = 1$, $x_2 = 0$ we have $\dot{x}_1 = q_{11}^{(1)} < 0$. Therefore $\dot{x}_1 = 0$ has a root on the boundary of E', and so it cannot correspond to an ellipse entirely in the interior of E'. Thus we now have the result that there can be at most four fixed points of the dynamical system.

Index theory

We continue to work in the two-dimensional state space E' defined in (4.9) with the vector field

$$\dot{x}_j = \sum_{i=1}^{3}\sum_{k=1}^{3} x_i x_k q_{ik}^{(j)}, \qquad j = 1, 2 \tag{4.11}$$

where the dummy variable x_3 is defined by $x_3 = 1 - x_1 - x_2$. Note that a consequence of lemma 4.1 is that the vector field points inwards on the boundary of E'. We will apply results from Guckenheimer & Holmes (1983) to deduce that there are an odd number of fixed points in E'. Suppose that C is a closed curve in E' which does not pass through any fixed points of the dynamical system. Then the index k of C is defined by

$$k = \frac{1}{2\pi}\int_C d\left\{\arctan\left(\frac{dx_1}{dx_2}\right)\right\}.$$

As we move around the curve C the index counts, positively in the anti-clockwise sense, the number of revolutions the direction of the vector field makes. The index of the curve defined by the boundary of E' is one, as the vector field always points inwards on E'. We can now apply the following result (see Guckenheimer & Holmes, 1983, Proposition 1.8.4):

(i) The index of a sink, a source or a center is $+1$.

(ii) The index of a hyperbolic saddle point is -1.

(iii) The index of a closed orbit is $+1$.

(iv) The index of a closed curve not containing any fixed points is 0.

(v) The index of a closed curve is equal to the sum of the indices of the fixed points within it.

Sinks, sources, centers and hyperbolic saddle points are all types of fixed points. This directly implies that we must have an odd number of fixed points in E'. Combining this with the previous result that there are at most four fixed points, we conclude that there are either one or three fixed points of the dynamical system.

The following two numerical examples show that both these cases can occur.

Example 1: Unique Fixed Point The Q-matrices are as follows:

$$\begin{aligned} Q^{(1)} &= \begin{pmatrix} -10 & 4 & 6 \\ 9 & -12 & 3 \\ 2 & 7 & -9 \end{pmatrix}, \\ Q^{(2)} &= \begin{pmatrix} -16 & 11 & 5 \\ 14 & -20 & 6 \\ 10 & 11 & -21 \end{pmatrix}, \\ Q^{(3)} &= \begin{pmatrix} -5 & 2 & 3 \\ 7 & -11 & 4 \\ 2 & 6 & -8 \end{pmatrix}. \end{aligned}$$

Numerical simulation of this dynamical system indicates that there is a unique fixed point at

$$x_1 = 0.4157389,\ x_2 = 0.3159193,\ x_3 = 0.2683418.$$

Independent of the choice of $x(0)$ we find that $x(t)$ tends to this fixed point as t tends to infinity.

Example 2: Three Fixed Points The Q-matrices are defined by

$$\begin{aligned}
Q^{(1)} &= \begin{pmatrix} -28 & 3 & 25 \\ 18 & -381 & 363 \\ 47 & 5 & -52 \end{pmatrix}, \\
Q^{(2)} &= \begin{pmatrix} -257 & 237 & 20 \\ 2 & -9 & 7 \\ 2 & 5 & -7 \end{pmatrix}, \\
Q^{(3)} &= \begin{pmatrix} -4 & 2 & 2 \\ 3 & -5 & 2 \\ 1 & 4 & -5 \end{pmatrix}.
\end{aligned}$$

Simulating this dynamical system from various starting points indicates that $x(t)$ tends to one of two fixed points as t tends to infinity. These fixed points are at

$$x_1 = 0.4056635,\ x_2 = 0.0542287,\ x_3 = 0.5401078$$

and at

$$x_1 = 0.0548061,\ x_2 = 0.2321006,\ x_3 = 0.7130933.$$

There is also an unstable fixed point at

$$x_1 = 0.0920227,\ x_2 = 0.1772681,\ x_3 = 0.7307093.$$

It seems harder to find examples where there are three fixed points. Relatively small changes to a couple of the elements of the Q-matrices in example 2 can result in a dynamical system with a single fixed point.

4.4 Discussion

We now return to the Markov process $X^N(t)$. Under assumption (4.2) it is clear that $X^N(t)$ is an irreducible recurrent finite state Markov process. Therefore there is a unique stationnary distribution μ^N for $X^N(t)$. Furthermore, as t tends to infinity the law of $X^N(t)$ tends to μ^N irrespective of $X^N(0)$.

In view of this, the results of section 4.3 are surprising. We proved that the sequence of processes $X^N(t)$ converge in distribution to X. However, the dynamical system X can have two stable fixed points. In such a case, $X(t)$ will tend to one of the fixed points as t tends to infinity. The starting point $X(0)$ will determine which fixed point the dynamical system tends to. Thus if the Q-matrices are such that the dynamical system has two fixed points, we can choose initial distributions $X^N(0)$ and $X(0)$ such that

$$X^N(0) \quad \Rightarrow \quad X(0)$$

but

$$\lim_{N\to\infty}\lim_{t\to\infty} X^N(t) \quad \not\Rightarrow \quad \lim_{t\to\infty} X(t).$$

(This is clear as the left-hand-side does not depend on the initial distributions but the right-hand-side does.) We do not have a contradiction here, however. The weak convergence that we have established states that

$$\lim_{N\to\infty} E\left[f(X^N)\right] \quad = \quad E\left[f(X)\right]$$

for every bounded continuous function $f : D_E[0,\infty) \to \mathbb{R}$. Convergence occurs under the Skorohod topology; functions f which capture the behaviour of an element of $D_E[0,\infty)$ in the limit as t tends to infinity will not be continuous in this topology.

The weak convergence does imply that

$$(X^N(t_1), X^N(t_2), \ldots, X^N(t_k)) \quad \Rightarrow \quad (X(t_1), X(t_2), \ldots, X(t_k))$$

for every k and finite set $\{t_1, t_2, \ldots, t_k\} \subset [0,\infty)$. From this we can conjecture the form of the stationary distributions μ^N of the processes X^N. We expect that the stationary distributions μ^N will have density increasingly concentrated around the fixed points of the dynamical system X as N increases. In cases such as example 1 of section 4.3.4 where the dynamical system X has a unique fixed point, the stationary distributions μ^N will will have density concentrated around the single fixed point. Where the dynamical system has more than one stable fixed point, as in example 2 of section 4.3.4, the stationary distributions μ^N will have density concentrated around each fixed point. For large t the process $X^N(t)$ will typically be close to a fixed point, occasionally moving from one fixed point to another. The larger the value of N, the less frequent these transitions are.

4.5 Conclusions

We began this chapter by observing that a strand of the economics literature has focused on search theoretic models in which there are a large number of agents, and pairs of these meet according to a random matching process. At these meeting times some event which changes the state of the agents may occur. We have analysed a Markov process which can be interpreted as a simple search model. Typically the state space of this Markov process is large; this makes numerical calculation of the stationary distribution difficult. We have attempted to learn about the Markov process through the dynamical system which is its limit as we let the number of agents increase.

A general result we have found is that there is always a fixed point of the dynamical system. In the relatively simple case of three possible states for each agent, we have proved that the dynamical system has either one or two stable fixed points. Examples show that both these cases can occur. The methods that we used to study the dynamical system with three states — conic sections and index theory — do not easily extend to four or more states. However we would expect that the maximum number of stable fixed points increases as the number of states for each agent increases.

Interesting questions remain concerning the stationary distribution of the Markov process. For example, where the corresponding dynamical system has multiple stable fixed points, we conjecture that the stationary distribution is concentrated near these fixed points. Future work could involve attempting to find how that stationary distribution is spread between the fixed points.

Chapter 5

Credit-risky Convertible Bonds

5.1 Introduction

A convertible bond is a financial instrument issued by a firm. As is the case with straight bonds, the owner of a convertible bond receives coupons from the firm up until the bond matures. However, the convertible feature means that under certain conditions the owner has the choice of converting the bond into shares in the firm. The owner would then receive the dividends paid to shareholders instead of the coupons. Some convertible bonds include call provision: if the firm chooses to call the convertibles, the bondholders must either convert their bonds into shares, or surrender their bonds in return for a final payment.

Two arguments are commonly put forward as to why firms issue convertible rather than straight debt (see Nyborg (1996)). The first is that convertible debt acts as "delayed equity", that is the issuing firm, expecting the debt to be converted at some point in the future, benefits from receiving payment immediately for equity that is not issued until conversion. The second is that the convertible feature acts as a "debt sweetener". The coupons on convertible debt can be set lower than those on straight debt, with the convertible feature compensating the buyer for this. From an investor's perspective, a convertible bond offers greater security than a share should the value of the firm fall. The investor can also benefit, through the convertible feature, should the value of the firm rise.

Ingersoll (1977) and Brennan & Schwartz (1977) apply the Black-Scholes methodology to the pricing of both callable and non-callable convertible debt. In both papers the value of the issuing firm is taken as the underlying exogenous process and is modelled as a diffusion process. The bondholders and shareholders make decisions based (at least in part) on the value of this process. Ingersoll (1977) concludes that, in his model, investors should never voluntarily convert their bonds except possibly when they mature. In the model of Brennan & Schwartz (1977), investors may also convert immediately prior to a dividend payment. However in both these papers should voluntary conversion occur, all bondholders will convert simultaneously. This feature seems undesirable.

In this chapter we take a firm-value approach to modelling the activity of a firm with an issue of non-callable perpetual convertible debt. We find the optimal behaviour for the management of the firm with regard to the timing of default, and the optimal behaviour for the bondholders with regard to voluntary conversion. This behaviour of the claim owners forms a Nash equilibrium. The prices of a share and of a convertible bond are thus found. Modelling perpetual convertible debt enables us to proceed further analytically than would be possible with finite maturity debt.

Leland (1994) and Leland & Toft (1996) study optimal conditions for default for a firm with an issue of straight debt. In these papers the value of the firm is modelled as a diffusion and the trigger for default is determined endogenously by the firm in order to maximise the value of the shares. We apply the same method in our model in order to find the optimal behaviour of the firm with regard to default.

We also permit bondholders to voluntarily convert at different times. The rationale behind this is that the voluntary conversion of a bond has two principal effects. Firstly the likelihood of bankruptcy is reduced, as the total coupon payments by the firm decrease. Secondly, the value of equity is diluted. If the effect of the latter is greater than that of the former, then it is conceivable that other bondholders will choose not to convert immediately. Emanuel (1983) and Constantinides (1984) look at the optimal exercise strategy for warrant holders who are not constrained to exercise their warrants simultaneously. Both authors conclude that simultaneous exercise of the warrants is suboptimal. As each author notes, this result also applies to convertible bonds. The framework that these papers work in does not include credit risk, however, and this result is therefore driven by the equity dilution effect.

In the next section we describe in detail the model.

5.2 The model

We consider a firm with equity, and a single issue of convertible debt. This convertible debt is assumed for simplicity to be perpetual, with each convertible bond attracting a continuous flow of coupons at rate ρ per unit time. The value V_t of the firm's assets at time t is modelled as a diffusion, satisfying

$$dV_t = V_t(\sigma dW_t + (r - \delta)dt), \tag{5.1}$$

where $r > 0$ is the riskless rate of interest, assumed constant, and $\delta > 0$, also assumed constant, represents the proportional rate of disbursement of value to the shareholders and bondholders, less the tax rebate received on the coupon payments. The model of Brennan & Schwartz (1980) permits stochastic interest rates. They observe that this increases the complexity of the problem significantly and appears to affect the results only slightly. We expect that the assumption of constant interest rates will have little qualitative effect on the behaviour of the claim owners. It is important to realise that δV_t is not the rate of payment of dividends to the shareholders: as in Leland (1994), and Leland & Toft (1996), we recognise that the provisions of corporate debt issues often explicitly prohibit the sale of the firm's assets in order to meet coupon payments. We take the tax rebate to be proportional to the coupon payments being made. Consequently, if there are m_t live convertible bonds at time t, the coupon payments are running at rate ρm_t and the firm receives the tax rebate at rate $\tau\rho m_t$ (for some constant τ between 0 and 1). Therefore the rate at which dividends are being paid to shareholders is

$$\delta V_t - (1-\tau)\rho m_t \equiv \delta V_t - \rho' m_t.$$

The dividend rate may of course be negative at times; if so, we interpret this as a call on the shareholders to inject more capital into the firm. The shareholders may be prepared to do this up to a point, but there may come a time when the rate of injection of capital required is so large that the shareholders would prefer to default, and throw the firm into bankruptcy. Should this happen, a proportion $1 - p$ of the firm's assets are lost, and what remains reverts to the bondholders.

A bondholder has the choice of when to convert his bond (if at all). We shall suppose that the terms of conversion are that the bondholder receives one new share for one bond. Thus we assume that the conversion ratio for the bonds is unity. Clearly there is no loss of generality in assuming this.

If there are originally N issued shares, and M convertible bonds, then should all the bonds be converted there will be $N + M$ shares; we shall let this number be denoted by n throughout the rest of this chapter. The state of the system at time t is uniquely described by the pair (m_t, V_t), where m_t is the number of live convertibles at time t, and we shall therefore have that the price S_t of a share at time t, and the price B_t of a convertible bond at time t will both be functions of (m_t, V_t) alone:

$$\begin{aligned} S_t &= S(m_t, V_t) \\ B_t &= B(m_t, V_t). \end{aligned}$$

Similarly, at time t the decision of the firm about whether to default, and the bondholders' decisions about whether to convert, will depend only on the pair (m_t, V_t).

We assume that the bonds are owned by a continuum of infinitesimal bondholders who do not collaborate with each other. This avoids the need to consider pre-emptive block exercise by different agents, and other game-theoretic aspects of bond conversion; see Constantinides & Rosenthal (1984) for a treatment of such issues in another example. By symmetry we expect that all bondholders will aim to follow the same policy in a Nash equilibrium.

We shall proceed on the assumption that bonds will be converted sequentially. Should several bondholders all wish to convert their bonds then the order in which conversion occurs is random. As bonds are converted the state of the system then changes and it may be that the decision of the remaining bondholders about whether to convert or not has changed. We will find that it is necessary for the analysis to specify that the final ε bonds can only be converted in one block. Thus the bonds will initially be converted sequentially, with the bondholders acting independently, until ε bonds remain (assuming that the firm does not default). At that point, the bondholders then act collectively in deciding whether to convert. The reason for this condition will be discussed in section 5.7. Typically ε will be small with respect to n. In fact, we require that ε is chosen small enough that

$$p/\varepsilon \geq 1/n. \tag{5.2}$$

To make the problem well-defined we must specify the result if the choices of the firm and those of the bondholders conflict. We will assume that the choice of the firm takes priority. For example, if simultaneously the firm chooses to default and a bondholder decides to convert his bond, then we assume that the firm defaults. This assumption has little bearing on the problem.

5.3 The solution

In this section we find the solution to the problem. The only decision for the firm is the timing of default, and the only decision for a bondholder is the timing of conversion. As discussed in the previous section, these decisions depend only on the pair (m, V). We assume that the objective of the firm is to maximise the value of the shares, and the objective for each bondholder is to maximise the value of his bond. What policies should the management of the firm and the bondholders follow?

The choice of policy for a bondholder will depend on how he expects the firm and the other bondholders to behave. Similarly, the firm, in choosing a policy, must make assumptions about the bondholders' policies. We aim to find functions S and B that correspond to the share and bond prices respectively, and policies Π_S and Π_B for the firm and bondholders respectively, which form a Nash equilibrium. This means that provided all agents follow their stated policy, no single agent can benefit from changing to a different policy. Functions S and B, and policies Π_S and Π_B which form a Nash equilibrium are a solution to the problem. We shall refer to these policies as optimal policies. Brennan & Schwartz (1977) define optimal policies in the same way in their context.

It is important to realise that when $m > \varepsilon$ the Nash equilibrium is made up of the firm acting as one agent, and each infinitesimal bondholder acting as an individual. The firm acts as one agent in deciding whether to default. Under our assumption that the ownership of the convertibles is diffuse, each bondholder can choose whether or not to convert his bond at any time, independently of the other bondholders. When $m = \varepsilon$ there are effectively two agents — the firm and the bondholders.

Theorem 5.1 claims that a solution exists and gives some properties of the optimal policies.

Theorem 5.1. *There exist* $m^* > \varepsilon$ *and functions* $\xi : [\varepsilon, n) \to \mathbb{R}^+$ *and* $\eta : [\varepsilon, n) \to \mathbb{R}^+$

satisfying the inequalities

$$\begin{aligned} \xi(m) &\leq \frac{m\rho'}{\delta} && (5.3) \\ \xi(m) &< \eta(m) && (5.4) \\ \frac{n\rho}{\delta} &\leq \eta(\varepsilon) && (5.5) \end{aligned}$$

such that

(i) η *is continuous on the interval* (ε, n),

(ii) ξ *is continuous,*

(iii) η *is constant on the interval* $(\varepsilon, m^*]$,

(iv) η *is decreasing, and strictly decreasing for* $m \geq m^*$,

in terms of which optimal policies are

Π_S: *the management of the firm defaults when* $V_t \leq \xi(m_t)$

Π_B: *the bondholders convert when* $V_t \geq \eta(m_t)$.

Thus the function $\xi(m)$ gives the default boundary, and the function $\eta(m)$ the conversion boundary. Figure 5-1 shows these boundaries in the (m, V) plane. We will find that typically m^* is close to ε and therefore small with respect to n. Figure 5-1 shows $\eta(\varepsilon)$ strictly greater than $\eta(m^*)$. It is possible that $\eta(\varepsilon) = \eta(m^*)$. The values of the parameters of the model determine which of these cases occurs.

We prove Theorem 5.1 by finding default and conversion boundaries, and the corresponding share and bond price functions which form a Nash equilibrium. The construction of $\xi(m)$ and $\eta(m)$ involves the functions S and B, defined below in terms of $\xi(m)$ and $\eta(m)$. We will find later that these functions give the share price and bond price.

The operator $\mathcal{L}$ is defined by

$$\mathcal{L} = \frac{\sigma^2}{2} V^2 \frac{\partial^2}{\partial V^2} + (r - \delta) V \frac{\partial}{\partial V} - r.$$

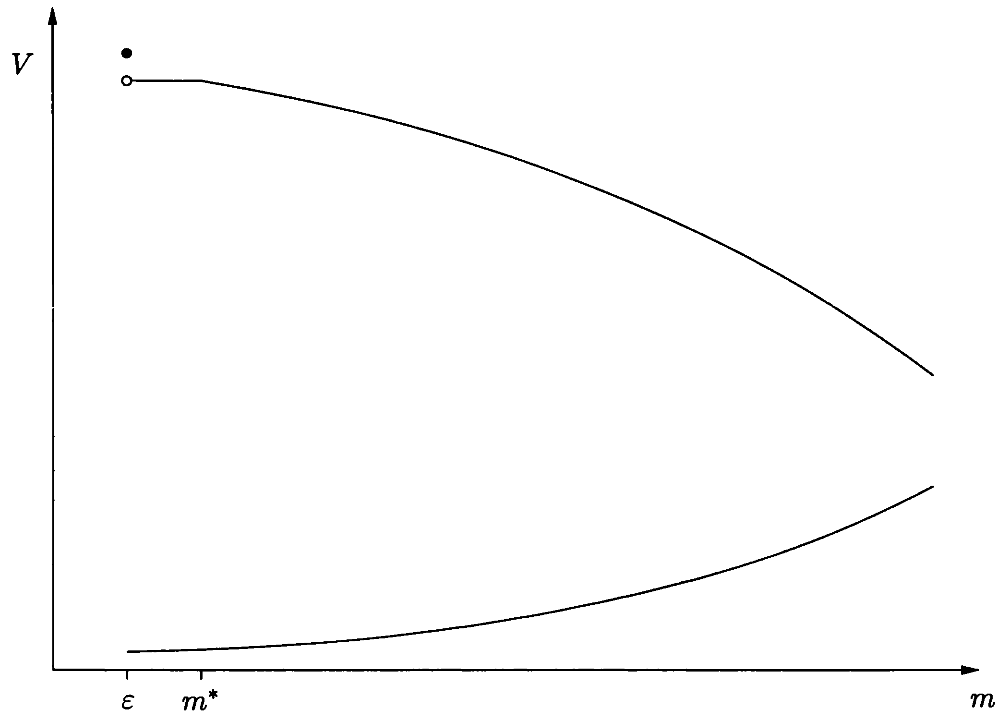

Figure 5-1: The boundaries $\xi(m)$ (lower curve) and $\eta(m)$ (upper curve) in the (m, V) plane.

We now define S and B.

Definition 5.1. *Given boundaries $\xi(m)$ and $\eta(m)$ which satisfy the inequalities (5.3)-(5.5) and conditions (i)-(iv) of Theorem 5.1, first construct $\tilde{S} : [\varepsilon, n) \times [0, \eta(\varepsilon)] \to \mathbb{R}$ by*

(S1) $\tilde{S}(m, V) = 0$ *for* $V \leq \xi(m)$

(S2) $\mathcal{L}\tilde{S} + \dfrac{\delta V - m\rho'}{n - m} = 0$ *for* $V \in [\xi(m), \eta(m)]$

(S3) $V \mapsto \tilde{S}(m, V)$ *is* C^1 *at* $\xi(m)$

and $\tilde{B} : [\varepsilon, n) \times [0, \eta(\varepsilon)] \to \mathbb{R}$ *by*

(B1) $\tilde{B}(m, V) = pV/m$ *for* $V \leq \xi(m)$

(B2) $\mathcal{L}\tilde{B} + \rho = 0$ *for* $V \in [\xi(m), \eta(m)]$

(B3) $\tilde{B}(m, \eta(m)) = \tilde{S}(m, \eta(m))$

(B4) $V \mapsto \tilde{B}(m,V)$ *is continuous at* $\xi(m)$ *and* $\eta(m)$.

Now define $\zeta : [0, \eta(\varepsilon)] \to [\varepsilon, n]$ *by*

$$\zeta(V) = \sup\{\varepsilon \le m < n : \eta(m) \ge V\}$$

and finally S *and* B *by*

$$\begin{aligned} S(m,V) &= \tilde{S}(m \wedge \zeta(V), V) && (5.6)\\ B(m,V) &= \tilde{B}(m \wedge \zeta(V), V) \quad \textit{for } 0 \le V \le \eta(m^*) && (5.7)\\ B(m,V) &= \frac{m-\varepsilon}{m}\tilde{S}(\varepsilon, V) + \frac{\varepsilon}{m}\tilde{B}(\varepsilon, V) \quad \textit{for } \eta(m^*) < V \le \eta(\varepsilon). && (5.8) \end{aligned}$$

Remarks

1. The solutions to the differential equations for S and B when $V \in (\xi(m), \eta(m))$ that arise from *(S2)* and *(B2)* are

$$\begin{aligned} S(m,V) &= \frac{V - m\rho'/r}{n-m} + a(m)V^{-\alpha} + b(m)V^{\beta} && (5.9)\\ B(m,V) &= \frac{\rho}{r} + a'(m)V^{-\alpha} + b'(m)V^{\beta} && (5.10) \end{aligned}$$

for any functions $a(m)$, $b(m)$, $a'(m)$ and $b'(m)$, where $-\alpha < 0$ and $\beta > 1$ are the roots to the quadratic

$$\frac{\sigma^2}{2}x(x-1) + (r-\delta)x - r = 0. \qquad (5.11)$$

This is easy to verify by direct differentiation. Observe that the definitions of α and β imply the following equalities:

$$\begin{aligned} \alpha\beta &= 2r/\sigma^2 && (5.12)\\ (\alpha+1)(\beta-1) &= 2\delta/\sigma^2. && (5.13) \end{aligned}$$

These will be used throughout this section.

2. Properties *(S1)* and *(S3)* imply that

$$S(m, \xi(m)) = 0$$

and

$$\frac{\partial S}{\partial V}(m, \xi(m)) \quad = \quad 0.$$

These two boundary conditions enable us to determine the functions $a(m)$ and $b(m)$ in (5.9) in terms of $\xi(m)$. We obtain

$$\begin{aligned} S(m,V) \quad = \quad & \frac{V - m\rho'/r}{n-m} + \frac{\beta m\rho'/r - (\beta-1)\xi(m)}{(\alpha+\beta)(n-m)}\left(\frac{V}{\xi(m)}\right)^{-\alpha} \\ & + \frac{\alpha m\rho'/r - (\alpha+1)\xi(m)}{(\alpha+\beta)(n-m)}\left(\frac{V}{\xi(m)}\right)^{\beta} \end{aligned} \quad (5.14)$$

for all $V \in (\xi(m), \eta(m))$. The continuity of ξ and η on (ε, n) implies that $S(m, \eta(m))$ is continuous on this interval; and from the strict monotonicity of η in (m^*, n) (see condition *(iv)* of Theorem 5.1) we deduce that the functions S and B defined above are continuous in $(m^*, n) \times [0, \eta(m^*)]$.

The continuity of $V \mapsto B(m, V)$ at $\xi(m)$ and $\eta(m)$ gives two boundary conditions for the differential equation for B arising from *(B2)*. We could determine $a'(m)$ and $b'(m)$ in (5.10) from these boundary conditions. The expression that we obtain is not useful for our analysis. However, it can be seen that knowledge of ξ and η allows us to construct S and B.

We now state and prove a theorem which gives sufficient properties of the functions S and B for them to correspond to the share price and bond price in a Nash equilibrium.

Theorem 5.2. *Suppose that we are given boundaries $\xi(m)$ and $\eta(m)$ which satisfy inequalities (5.3)-(5.5) and conditions (i)-(iv) of Theorem 5.1. If the functions S and B of definition 5.1 have the properties*

(P1) $S(\varepsilon, \eta(\varepsilon)) = \eta(\varepsilon)/n = B(\varepsilon, \eta(\varepsilon))$

(P2) $S(\varepsilon, \eta(m^*)) = S(m, \eta(m^*))$ *for* $\varepsilon \le m \le m^*$

then we extend S and B continuously to $[\varepsilon, n) \times \mathbb{R}^+$ by

(P3) $B(m, V) = S(m, V) = V/n$ *for* $m \in [\varepsilon, n)$ *and* $V \ge \eta(\varepsilon)$.

Suppose that the functions S and B so extended have the properties

(P4) $V \mapsto B(\varepsilon, V)$ *is* C^1 *at* $\eta(\varepsilon)$

(P5) $B(\varepsilon, \eta(m^*)) = S(\varepsilon, \eta(m^*))$

(P6) the left-derivative of $V \mapsto S(m,V) - B(m,V)$ *is zero at* $\eta(m)$ *for* $m \geq m^*$

(P7) $B(\varepsilon, V) \geq V/n$

(P8) $S(m,V) < B(m,V)$ *for* $m > \varepsilon$, $V \in (\xi(m), \eta(m))$

(P9) $S(m,V) \geq 0$ *for* $m < \bar{m}$

(P10) $B(\varepsilon, V) \leq S(\varepsilon, V)$ *for* $V \in (\eta(m^*), \eta(\varepsilon))$

(P11) for all m*, the partial derivatives* $\dfrac{\partial S}{\partial m}(m, \eta(m))$ *and* $\dfrac{\partial B}{\partial m}(m, \eta(m))$ *exist and are zero*

for some $\bar{m}$*. Then* S *corresponds to the share price,* B *corresponds to the bond price and the policies* Π_S *and* Π_B *are optimal for* $m < \bar{m}$*.*

Remarks

1. A consequence of *(P2)* and condition (*iii*) of Theorem 5.1 is that $S(m, \eta(m))$ must be constant for m in the interval $[\varepsilon, m^*]$. Together with *(B3)* and *(P5)* this implies that

$$B(m, \eta(m)) = S(m, \eta(m)) = S(\varepsilon, \eta(\varepsilon)) \qquad \text{for } m \in [\varepsilon, m^*]. \tag{5.15}$$

2. We will construct boundaries ξ and η, and so functions S and B, such that we can apply Theorem 5.2 with $\bar{m} = n$. This is sufficient to prove Theorem 5.1. We are unable to prove that the boundary η which we construct is strictly decreasing on (m^*, n), as required by condition (*iv*) of Theorem 5.1. This is left as an assumption and is supported by numerical evidence. For further discussion of this see the comments following the proof of lemma 5.4 and section 5.7.

Proof of Theorem 5.2. Suppose that $m_0 < \bar{m}$. Define X_t by

$$\begin{aligned} X_t &= e^{-rt}S(m_t, V_t) - \int_0^t e^{-ru}\mathcal{L}S(m_u, V_u)du - \int_0^t e^{-ru}\frac{\partial S}{\partial m}(m_{u-}, V_{u-})dm_u \\ &\quad - \sum_{0<u\le t} e^{-ru}\left\{S(m_u, V_u) - S(m_{u-}, V_{u-}) - \frac{\partial S}{\partial m}(m_{u-}, V_{u-})\Delta m_u\right\} \qquad (5.16) \end{aligned}$$

and define $S(0, V)$ to be the share price when no bonds remain, that is

$$S(0, V) = \frac{V}{n}.$$

Itô's formula implies that X_t is a local martingale[1]. Therefore for any stopping time T which reduces X_t we have

$$\begin{aligned} X_0 &= S(m_0, V_0) = E[X_T] \\ &= E\Bigg[e^{-rT}S(m_T, V_T) - \int_0^T e^{-ru}\mathcal{L}S(m_u, V_u)du - \int_0^T e^{-ru}\frac{\partial S}{\partial m}(m_{u-}, V_{u-})dm_u \\ &\quad - \sum_{0<u\le T} e^{-ru}\left\{S(m_u, V_u) - S(m_{u-}, V_{u-}) - \frac{\partial S}{\partial m}(m_{u-}, V_{u-})\Delta m_u\right\}\Bigg] \qquad (5.17) \end{aligned}$$

Now suppose that the bondholders follow policy Π_B. We aim to show that Π_S is the optimal policy for the firm. Under Π_B, if $V_0 \le \eta(m^*)$ then the bondholders convert continuously except for two jumps, from $m = m^*$ down to $m = \varepsilon$ and from $m = \varepsilon$ down to $m = 0$[2]. If $V_0 > \eta(m^*)$ then there is also a jump at time 0. Since $\frac{\partial S}{\partial m} = 0$ along the continuous conversion boundary by property *(P11)*, and there is no change in S when any jumps of m occur, the terms involving changes in m in (5.17) all disappear for any policy that the firm follows. We are left with

$$S(m_0, V_0) = E\left[e^{-rT}S(m_T, V_T) - \int_0^T e^{-ru}\mathcal{L}S du\right] \qquad (5.18)$$

for any stopping time T which strongly reduces X.

[1]The function $V \mapsto S(m, V)$ is not C^2 at $\xi(m)$. However it is C^1 (property *(S3)*) and the second derivative exists as a measure. Itô's formula can be extended to deal with this. If S were not C^1 at ξ, there would be an additional term in (5.16) involving the local time of V at ξ, whose effect would be to reduce the expression, since *(S1)* and *(P9)* imply that $\frac{\partial S}{\partial V}(m, \xi(m)-) = 0 \le \frac{\partial S}{\partial V}(m, \xi(m)+)$.

[2]If $\eta(m^*) = \eta(\varepsilon)$ then these two jumps coincide and become a single jump. It is valid to consider this case as one jump immediately followed by another.

Define

$$H_0 \quad = \quad \inf\{t : m_t = 0\} \tag{5.19}$$

and suppose that T_n are stopping times increasing to infinity that reduce X. Suppose also that the firm chooses the rule 'default at $\bar{T}$' for some stopping time $\bar{T}$. I claim that

$$\mathcal{L}S + \frac{\delta V - m\rho'}{n - m} \;\leq\; 0 \tag{5.20}$$

in the region of the (m, V) plane defined by the policy Π_B. This holds by construction for $V \in (\xi(m), \eta(m))$. When $V < \xi(m)$ we have from *(S1)* that $\mathcal{L}S(m, V) = 0$ and so the bound on $\xi(m)$ in (5.3) is enough to give (5.20). From (5.18) we obtain

$$\begin{aligned}
&S(m_0, V_0) \\
\geq\;& E\left[e^{-r(T_n \wedge H_0 \wedge \bar{T})} S(m_{T_n \wedge H_0 \wedge \bar{T}}, V_{T_n \wedge H_0 \wedge \bar{T}}) + \int_0^{T_n \wedge H_0 \wedge \bar{T}} e^{-ru} \frac{\delta V_u - m_u \rho'}{n - m_u} du\right] \\
=\;& E\left[e^{-r(H_0 \wedge \bar{T})} S(m_{H_0 \wedge \bar{T}}, V_{H_0 \wedge \bar{T}}) + \int_0^{H_0 \wedge \bar{T}} e^{-ru} \frac{\delta V_u - m_u \rho'}{n - m_u} du; T_n \geq H_0 \wedge \bar{T}\right] \\
&+ E\left[e^{-rT_n} S(m_{T_n}, V_{T_n}) + \int_0^{T_n} e^{-ru} \frac{\delta V_u - m_u \rho'}{n - m_u} du; T_n < H_0 \wedge \bar{T}\right]
\end{aligned} \tag{5.21}$$

Now consider what happens to the two terms in (5.21) as n increases to infinity. As V is bounded on the time interval $[0, H_0]$ the second term will tend to zero. The first term will tend to

$$\begin{aligned}
&E\left[e^{-r(H_0 \wedge \bar{T})} S(m_{H_0 \wedge \bar{T}}, V_{H_0 \wedge \bar{T}}) + \int_0^{H_0 \wedge \bar{T}} e^{-ru} \frac{\delta V_u - m_u \rho'}{n - m_u} du\right] \\
\geq\;& E\left[e^{-r(H_0 \wedge \bar{T})} \frac{V_{H_0}}{n} I\left\{H_0 < \bar{T}\right\} + \int_0^{H_0 \wedge \bar{T}} e^{-ru} \frac{\delta V_u - m_u \rho'}{n - m_u} du\right]
\end{aligned}$$

where the inequality follows from *(P9)*. Thus $S(m_0, V_0)$ is an upper bound for the net present value of the future cashflow from a share under any default policy. However, if the management of the firm follows the policy Π_S and defaults when $V_t \leq \xi(m_t)$ then both inequalities become equalities, the first because *(S2)* shows that (5.20) is an equality for $\xi(m) < V < \eta(m)$ and the second because *(S1)* shows that $S(m_{\bar{T}}, V_{\bar{T}}) = 0$ on $\{\bar{T} < H_0\}$.

We now apply a similar argument to the function B. We will assume that all bondholders follow the policy Π_B and deduce that the function $B(m, V)$ gives the value of the bond, that is the net present value of all future cashflows. First we will assume that we are considering a bondholder who is one of the last to convert — he converts at $m = \varepsilon$ (provided the firm does not default before then) — and that $V_0 \leq \eta(m_0)$. Once we have considered this case the results for the bondholders who convert earlier under Π_B, and for the case where $V_0 > \eta(m_0)$ will follow. We retain the assumption that $m_0 < \bar{m}$.

Let $\bar{T}$ denote the time that the management of the firm defaults under Π_S. Define Z_t by

$$\begin{aligned} Z_t &= e^{-r(t\wedge\bar{T})}B(m_{t\wedge\bar{T}}, V_{t\wedge\bar{T}}) - \int_0^{t\wedge\bar{T}} e^{-ru}\mathcal{L}B(m_u, V_u)du \\ &\quad - \int_0^{t\wedge\bar{T}} e^{-ru}\frac{\partial B}{\partial m}(m_u, V_u)dm_u \\ &\quad - \sum_{0<u\leq t\wedge\bar{T}} e^{-ru}\left\{B(m_u, V_u) - B(m_{u-}, V_{u-}) - \frac{\partial B}{\partial m}(m_{u-}, V_{u-})\Delta m_u\right\} \end{aligned} \tag{5.22}$$

This is a local martingale as a consequence of Itô's formula. Therefore if T reduces Z we have

$$\begin{aligned} Z_0 &= B(m_0, V_0) = E[Z_T] \\ &= E\Bigg[e^{-r(T\wedge\bar{T})}B(m_{T\wedge\bar{T}}, V_{T\wedge\bar{T}}) - \int_0^{T\wedge\bar{T}} e^{-ru}\mathcal{L}B(m_u, V_u)du \\ &\quad - \int_0^{T\wedge\bar{T}} e^{-ru}\frac{\partial B}{\partial m}(m_u, V_u)dm_u \\ &\quad - \sum_{0<u\leq T\wedge\bar{T}} e^{-ru}\left\{B(m_u, V_u) - B(m_{u-}, V_{u-}) - \frac{\partial B}{\partial m}(m_{u-}, V_{u-})\Delta m_u\right\}\Bigg] \end{aligned}$$

Now the conversion of bonds will be continuous except for the two jumps, from $m = m^*$ to $m = \varepsilon$ and from $m = \varepsilon$ to $m = 0$[3]. We have found that B remains constant when either of these jumps occurs (see (5.15) and *(P3)*). Since $\frac{\partial B}{\partial m} = 0$ as the bonds are continuously converted (property *(P11)*) we deduce that

$$B(m_0, V_0) = E\left[e^{-r(T\wedge\bar{T})}B(m_{T\wedge\bar{T}}, V_{T\wedge\bar{T}}) - \int_0^{T\wedge\bar{T}} e^{-ru}\mathcal{L}B(m_u, V_u)du\right]$$

[3]Again, it does not matter that these two jumps may coincide and become a single jump.

for any stopping time T which strongly reduces Z.

Recall from (5.19) the definition of H_0 and let T_n be a sequence of stopping times increasing to infinity which reduce Z. In the region of the (m, V) plane defined by the policies Π_S and Π_B we have, from $(B2)$, that

$$\mathcal{L}B + \rho = 0$$

and hence

$$\begin{aligned}
B(m_0, V_0) &= E\left[e^{-r(T_n \wedge H_0 \wedge \bar{T})} B(m_{T_n \wedge H_0 \wedge \bar{T}}, V_{T_n \wedge H_0 \wedge \bar{T}}) + \int_0^{T_n \wedge H_0 \wedge \bar{T}} e^{-ru} \rho du\right] \\
&= E\left[e^{-r(H_0 \wedge \bar{T})} B(m_{H_0 \wedge \bar{T}}, V_{H_0 \wedge \bar{T}}) + \int_0^{H_0 \wedge \bar{T}} e^{-ru} \rho du \; ; \; T_n \geq H_0 \wedge \bar{T}\right] \\
&\quad + E\left[e^{-rT_n} B(m_{T_n}, V_{T_n}) + \int_0^{T_n} e^{-ru} \rho du \; ; \; T_n < H_0 \wedge \bar{T}\right].
\end{aligned}$$

Similarly to before, the second term here tends to zero as n increases to infinity. The first term tends to

$$\begin{aligned}
&E\left[e^{-rH_0} B(0, V_{H_0}) + \int_0^{H_0} e^{-ru} \rho du \; ; \; H_0 < \bar{T}\right] \\
&+ E\left[e^{-r\bar{T}} \frac{pV_{\bar{T}}}{m_{\bar{T}}} + \int_0^{\bar{T}} e^{-ru} \rho du \; ; \; \bar{T} < H_0\right].
\end{aligned}$$

If we interpret $B(0, V_{H_0})$ as the value of a share at $m = 0$ and $V = V_{H_0}$ (the owner of the bond will have just converted the bond to a share) then we see that $B(m_0, V_0)$ equals the net present value of the future cashflow to the owner of the bond.

Now consider a bondholder who converts when $m > \varepsilon$ under Π_B. By the definition of Π_B this bondholder converts at $V = \eta(m)$. Property *(B3)* shows that the functions S and B are equal here. Therefore the net present value of the future cashflows for a shareholder and a bondholder who does not end up converting until $m = \varepsilon$ are equal. The converting bondholder is swapping one cashflow for another with the same net present value. Thus $B(m_0, V_0)$ equals the net present value of the future cashflows to bondholders who convert before $m = \varepsilon$ under Π_B, as well as those who convert at $m = \varepsilon$ under Π_B.

Next we aim to prove that any alternative policy for a bondholder results in a decrease in the net present value of his future cashflows. We assume that the management of the firm follows Π_S and the other bondholders follow Π_B. Let H_ε be defined by

$$H_\varepsilon = \inf\{t : m_t = \varepsilon\}.$$

The process V_t remains on the interval $[\xi(m_t), \eta(m_t)]$ for $t < H_\varepsilon$. If a bondholder follows an alternative policy and ends up converting before H_ε then we must have $V_t < \eta(m_t)$ at the time of conversion. As we have identified the functions S and B with the share and bond prices under Π_S and Π_B, property *(P8)* shows that this bondholder has decreased the net value of his future cashflows. Thus his behaviour is suboptimal.

The case where ε bondholders remain and act collectively needs to be treated separately. Suppose that $m_0 = \varepsilon$ and recall the definition of $\bar{T}$ as the time of default under Π_S, and of the local martingale Z_t in (5.22). A policy for the bondholders now corresponds to a choice of H_0, defined in (5.19). We have[4]

$$\begin{aligned}
Z_0 &= B(m_0, V_0) = E[Z_{H_0}] \\
&= E\Bigg[e^{-r(H_0\wedge\bar{T})}B(m_{H_0\wedge\bar{T}}, V_{H_0\wedge\bar{T}}) - \int_0^{H_0\wedge\bar{T}} e^{-ru}\mathcal{L}B(m_u, V_u)du \\
&\quad - \int_0^{H_0\wedge\bar{T}} e^{-ru}\frac{\partial B}{\partial m}(m_u, V_u)dm_u \\
&\quad - \sum_{0<u\le H_0\wedge\bar{T}} e^{-ru}\left\{B(m_u, V_u) - B(m_{u-}, V_{u-}) - \frac{\partial B}{\partial m}(m_{u-}, V_{u-})\Delta m_u\right\}\Bigg]
\end{aligned}$$

Property *(P7)* shows that $B(m, V)$ can only decrease at the point of conversion[5]. For $V \in (\xi(\varepsilon), \eta(\varepsilon))$ we have from *(B2)* that

$$\mathcal{L}B(\varepsilon, V) = -\rho.$$

For $V > \eta(m)$, property *(P7)* implies that

$$\mathcal{L}B(\varepsilon, V) = -\frac{\delta V}{n}.$$

[4]If we did not have the C^1 condition on $V \mapsto B(\varepsilon, V)$ at $\eta(\varepsilon)$ of *(P4)* then there would be an extra term in this expression for Z_0 involving the local time of V at $\eta(\varepsilon)$. Since *(P3)* and *(P7)* imply that $\frac{\partial B}{\partial V}(\varepsilon, \eta(\varepsilon)-) \le 1/n = \frac{\partial B}{\partial V}(\varepsilon, \eta(\varepsilon)+)$ this term would decrease Z_0.

[5]Again $B(0, V)$ is interpreted as the value a share at $m = 0$.

The bound on $\eta(\varepsilon)$ in (5.5) then shows that $-\rho$ is an upper bound for $\mathcal{L}B(\varepsilon, V)$. Thus

$$B(m_0, V_0) \quad \geq \quad E\left[e^{-r(H_0\wedge\bar{T})}B(m_{H_0\wedge\bar{T}}, V_{H_0\wedge\bar{T}}) + \int_0^{H_0\wedge\bar{T}} e^{-ru}\rho du\right]$$

and it follows that no alternative policy can improve upon Π_B.

Finally consider the situation if $V_0 > \eta(m_0)$. Under Π_B there will be immediate conversion and an immediate jump in m_t. We consider three cases here.

If $m_0 > m^*$ and $\eta(m_0) < V_0 \leq \eta(m^*)$ then (5.6), (5.7) and *(B3)* show that the value of the bond and share are equal and do not change as the jump in m occurs. Thus $B(m_0, V_0)$ corresponds to the net time-0 value of the future cashflows to a bondholder. Furthermore a bondholder who follows an alternative policy and chooses not to convert does not benefit — as the other bondholders convert the value of the bond remains constant.

If $\eta(m^*) < V_0 < \eta(\varepsilon)$ then the bonds are converted instantly down to $m = \varepsilon$. Under Π_B all bondholders will attempt to convert; the order of conversion is chosen at random. Thus an infinitesimal bondholder will convert with probability $\frac{m-\varepsilon}{m}$ and will remain a bondholder with probability $\frac{\varepsilon}{m}$. Therefore (5.8) shows that $B(m_0, V_0)$ corresponds to the net time-0 value of the future cashflows from the bond. An alternative policy for a bondholder is not to convert. By doing this the bondholder ensures that he holds a bond when $m - \varepsilon$ other bonds have been converted. Now (5.8) and *(P10)* show that $B(m_0, V_0)$ becomes an upper bound for the net time-0 value of the cashflows to the bondholder under this alternative policy. Therefore this behaviour is suboptimal.

The last case is when $V_0 \geq \eta(\varepsilon)$. Property *(P3)* shows that

$$S(m, V_0) = B(m, V_0) = \frac{V}{n}$$

for all m. We have already shown that it is suboptimal for the final ε bondholders to delay conversion if $V > \eta(\varepsilon)$. This case is therefore straightforward as all bonds are converted immediately. $\square$

It remains to define the boundaries $\xi(m)$ and $\eta(m)$ and to prove that S and B have the required properties. From (5.9) and (5.10) we see that the functions S and B have a similar form — the sum of terms linear in V, a term in $V^{-\alpha}$ and a term in V^{β}. We will find throughout this section that we are dealing with functions of this form. It is helpful now to present a lemma which establishes some properties of such functions.

Lemma 5.1. *If $f : (0, \infty) \to \mathbb{R}$ is of the form*

$$f(x) = a + bx + cx^{-\alpha} + dx^{\beta}$$

for constants a, b, c, d, with c and d not both equal to zero then f'' has at most one zero and f' at most two zeros. Furthermore, if $b \geq 0$ then f cannot have a maximum and then a minimum; if $b \leq 0$ then f cannot have a minimum and then a maximum.

Proof of lemma 5.1. The second derivative of f is

$$f''(x) = \alpha(\alpha+1)cx^{-\alpha-2} + \beta(\beta-1)dx^{\beta-2}$$

which has at most one root. Therefore f' has at most two zeros. Suppose $b \geq 0$ and f has a maximum and then a minimum. These would then be the only two zeros of f', which together with the form of f would imply that

$$\lim_{x \downarrow 0} f(x) = -\infty$$

and

$$\lim_{x \uparrow \infty} f(x) = \infty.$$

This is only possible if $c < 0$ and $d > 0$. However if this were the case then we would have

$$\begin{aligned} f'(x) &= b - \alpha c x^{-\alpha-1} + \beta d x^{\beta-1} \\ &> 0 \end{aligned}$$

which is a contradiction, since f' has two zeros. A similar argument shows that if $b \leq 0$ then f cannot have a minimum and then a maximum. $\square$

We now proceed to define the boundaries $\xi(m)$ and $\eta(m)$. We begin by defining the values of $\xi(\varepsilon)$ and $\eta(\varepsilon)$. The form of $S(\varepsilon, V)$ for $V \in (\xi(\varepsilon), \eta(\varepsilon))$ is given in (5.14). We must choose $\xi(\varepsilon)$ and $\eta(\varepsilon)$ so that $V \mapsto S(\varepsilon, V)$ is continuous at $\eta(\varepsilon)$.

The continuity of $V \mapsto B(\varepsilon, V)$ at $\eta(m)$ and property *(P4)* imply that

$$B(\varepsilon, \eta(\varepsilon)) \quad = \quad \frac{\eta(\varepsilon)}{n}$$

and

$$\frac{\partial B}{\partial V}(\varepsilon, \eta(\varepsilon)) \quad = \quad \frac{1}{n}.$$

Thus we can solve for the functions $a'(\varepsilon)$ and $b'(\varepsilon)$ in (5.10) to obtain

$$\begin{aligned} B(\varepsilon, V) \quad = \quad & \frac{\rho}{r} - \frac{\beta\rho/r - (\beta - 1)\eta(\varepsilon)/n}{\alpha + \beta}\left(\frac{V}{\eta(\varepsilon)}\right)^{-\alpha} \\ & - \frac{\alpha\rho/r - (\alpha + 1)\eta(\varepsilon)/n}{\alpha + \beta}\left(\frac{V}{\eta(\varepsilon)}\right)^{\beta}. \end{aligned} \tag{5.23}$$

for $V \in (\xi(\varepsilon), \eta(\varepsilon))$. This defines $B(\varepsilon, V)$ on this interval. The choice of $\xi(\varepsilon)$ and $\eta(\varepsilon)$ must ensure that $V \mapsto B(\varepsilon, V)$ is continuous at $\xi(\varepsilon)$. Lemma 5.2 proves that a solution exists.

Lemma 5.2. *There exist $\xi(\varepsilon)$ and $\eta(\varepsilon)$ satisfying the bounds in (5.3)-(5.5) such that the functions $V \mapsto S(\varepsilon, V)$ and $V \mapsto B(\varepsilon, V)$ are continuous and*

$$B(\varepsilon, V) \quad \geq \quad V/n. \tag{5.24}$$

Figures 5-2 and 5-3 show examples of the functions $S(\varepsilon, V)$ and $B(\varepsilon, V)$ that lemma 5.2 describes. In figure 5-2 the function $S(\varepsilon, V)$ crosses the line V/n between $\xi(\varepsilon)$ and $\eta(\varepsilon)$ and then meets this line from above at $V = \eta(\varepsilon)$. In figure 5-3 the function $S(\varepsilon, V)$ remains below the line V/n between $\xi(\varepsilon)$ and $\eta(\varepsilon)$ and approaches this line from below at $V = \eta(\varepsilon)$. The particular parameter values chosen determine which one of these figures is appropriate.

Proof of lemma 5.2. I claim that for any candidate value of $\eta(\varepsilon)$ there exists a unique $\xi(\varepsilon)$ on the interval

$$(0, \min\{\eta(\varepsilon), \varepsilon\rho'/\delta\})$$

such that the function $S(\varepsilon, V)$ is continuous in V at $V = \eta(\varepsilon)$. Recall from *(S5)* that $S(\varepsilon, \eta(\varepsilon)) = \eta(\varepsilon)/n$.

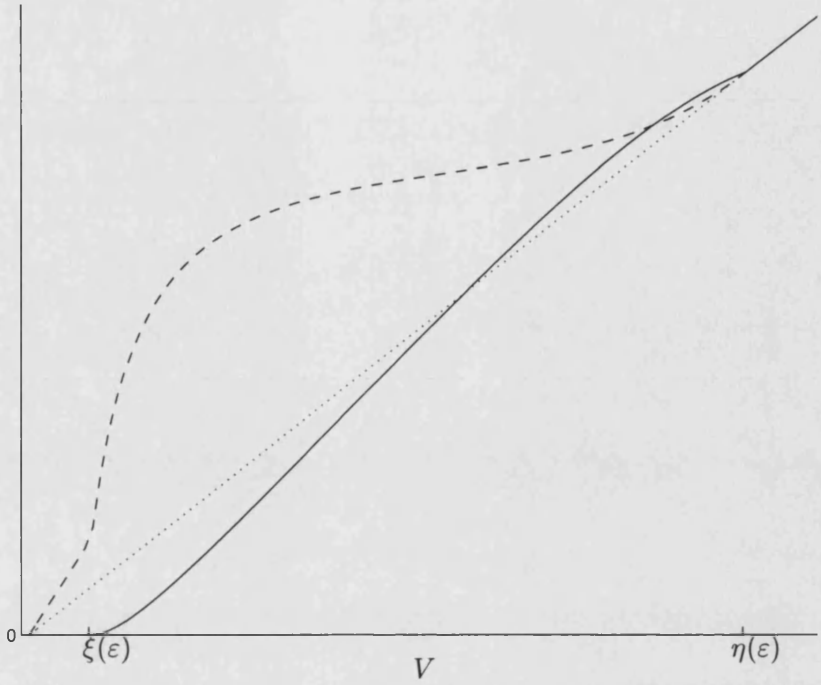

Figure 5-2: Possible form of $S(\varepsilon, V)$ (solid line) and $B(\varepsilon, V)$ (dashed line) with these functions crossing below the conversion boundary. The dotted line is V/n.

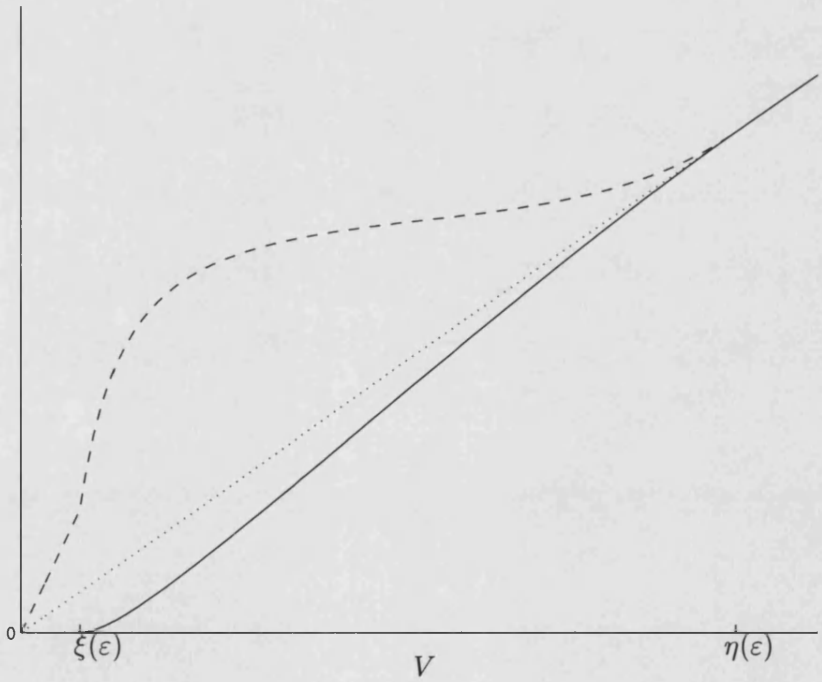

Figure 5-3: Possible form of $S(\varepsilon, V)$ (solid line) and $B(\varepsilon, V)$ (dashed line) with these functions not meeting until the conversion boundary. The dotted line is V/n.

To see this first consider the expression in (5.14) for $S(\varepsilon, V)$ as a function of ξ for fixed V. That is, consider the expression

$$\begin{aligned} \frac{V-\varepsilon\rho'/r}{n-\varepsilon} &+ \frac{\beta\varepsilon\rho'/r-(\beta-1)\xi}{(\alpha+\beta)(n-\varepsilon)}\left(\frac{V}{\xi}\right)^{-\alpha} \\ &+ \frac{\alpha\varepsilon\rho'/r-(\alpha+1)\xi}{(\alpha+\beta)(n-\varepsilon)}\left(\frac{V}{\xi}\right)^{\beta} \end{aligned}$$

for fixed V. It is apparent that as ξ decreases to zero, this expression tends to infinity. If we choose a small enough $\xi(\varepsilon)$, therefore, $S(\varepsilon, V)$ will have a negative jump at $V = \eta(\varepsilon)$.

The expression for $S(\varepsilon, V)$ given in (5.14) equals zero and has zero derivative at $V = \xi(\varepsilon)$. So if $\eta(\varepsilon) \le \varepsilon\rho'/\delta$, choosing $\xi(\varepsilon)$ close enough to but less than $\eta(\varepsilon)$ will result in $S(\varepsilon, V)$ having a positive jump at $V = \eta(\varepsilon)$. If $\eta(\varepsilon) > \varepsilon\rho'/\delta$ then choosing $\xi(\varepsilon) = \varepsilon\rho'/\delta$ yields

$$(\alpha+\beta)(n-\varepsilon)V^2\frac{\partial^2 S}{\partial V^2}(\varepsilon, V) = \frac{2\varepsilon\rho'}{\sigma^2}\left[\left(\frac{V}{\xi}\right)^{-\alpha} - \left(\frac{V}{\xi}\right)^{\beta}\right]$$

for $\xi(\varepsilon) < V < \eta(\varepsilon)$. This equation can be obtained by differentiating (5.14) and using the identities (5.12) and (5.13). The right-hand-side is negative for $V > \xi(\varepsilon)$ and so $V \mapsto S(\varepsilon, V)$ is concave for $V \in (\xi(\varepsilon), \eta(\varepsilon))$. As $V \mapsto S(\varepsilon, V)$ equals zero and has zero derivative at $V = \xi(\varepsilon)$, we conclude that $S(\varepsilon, V) < 0$ for V on this interval. Thus in this case also $S(\varepsilon, V)$ has a positive jump at $V = \eta(\varepsilon)$.

The expression for $S(\varepsilon, V)$ on the interval $(\xi(\varepsilon), \eta(\varepsilon))$ is continuous in $\xi(\varepsilon)$. Therefore we deduce that there exists a value of $\xi(\varepsilon)$ satisfying the bounds in (5.3) which makes the definition of $S(\varepsilon, V)$ continuous for any candidate $\eta(\varepsilon)$.

Next we will show that this $\xi(\varepsilon)$ is unique. Suppose there were two such values of $\xi(\varepsilon)$. Then consider the two expressions for $S(\varepsilon, V)$ in the interval $(\xi(\varepsilon), \eta(\varepsilon))$. Lemma 5.1 implies that both expressions have minima at their respective values of $\xi(\varepsilon)$. As these functions cannot have a maximum and then a minimum, both must tend to infinity as $V \downarrow 0$. Therefore there exists some V between the two values of $\xi(\varepsilon)$ at which the expressions are equal. We also know, by construction, that the expressions are equal at $V = \eta(\varepsilon)$. This gives the required contradiction as the form of the difference is a non-zero sum of a multiple of $V^{-\alpha}$ and a multiple of V^{β} which can have at most one root.

The next step is to show that we can choose $\eta(\varepsilon)$ so that $B(\varepsilon, V)$ is continuous at $V = \xi(\varepsilon)$. First suppose our candidate $\eta(\varepsilon)$ is less than $n\rho/\delta$. Then we find the $\xi(\varepsilon)$ which makes $S(\varepsilon, V)$ continuous at $V = \eta(\varepsilon)$. The definition of $B(\varepsilon, V)$ ensures that it is C^1 at $V = \eta(\varepsilon)$. By differentiating (5.23) and applying identities (5.12) and (5.13) we find that, for $V \in (\xi(\varepsilon), \eta(\varepsilon))$,

$$\begin{aligned}
\frac{\sigma^2}{2}(\alpha+\beta)V^2\frac{\partial^2 B}{\partial V^2}(\varepsilon, V) &= -((\alpha+1)\rho - \alpha\delta\eta(\varepsilon)/n)\left(\frac{V}{\eta(\varepsilon)}\right)^{-\alpha} \\
&\quad -((\beta-1)\rho - \beta\delta\eta(\varepsilon)/n)\left(\frac{V}{\eta(\varepsilon)}\right)^{\beta} \\
&< \rho\left\{\left(\frac{V}{\eta(\varepsilon)}\right)^{\beta} - \left(\frac{V}{\eta(\varepsilon)}\right)^{-\alpha}\right\} \\
&< 0.
\end{aligned}$$

Therefore $B(\varepsilon, V)$ is concave in V for $V \in (\xi(\varepsilon), \eta(\varepsilon))$, and so is less than V/n for all $V \in (\xi(\varepsilon), \eta(\varepsilon))$. This implies that there is a negative discontinuity in B at $V = \xi(\varepsilon)$.

Now we show that by choosing $\eta(\varepsilon)$ large enough there is a positive discontinuity in B at $V = \xi(\varepsilon)$. We have already found that, given $\eta(\varepsilon)$, the $\xi(\varepsilon)$ which makes $S(\varepsilon, V)$ continuous at $V = \eta(\varepsilon)$ is bounded above by $\varepsilon\rho'/\delta$. Consider the expression for $B(\varepsilon, V)$ in (5.23). For large enough $\eta(\varepsilon)$, the coefficient of V^β in (5.23) is positive. Therefore we have

$$B(\varepsilon, V) > \frac{\rho}{r} - \frac{\beta\rho/r - (\beta-1)\eta(\varepsilon)/n}{\alpha+\beta}\left(\frac{V}{\eta(\varepsilon)}\right)^{-\alpha} \tag{5.25}$$

for large enough $\eta(\varepsilon)$. As $\eta(\varepsilon)$ tends to infinity the right-hand-side of (5.25) also tends to infinity uniformly in V for $V \in (0, \varepsilon\rho'/\delta)$. Thus, as the expression for $B(\varepsilon, V)$ on the interval $V \in (\xi(\varepsilon), \eta(\varepsilon))$ is continuous in $\eta(\varepsilon)$, there exists a value of $\eta(\varepsilon)$ satisfying (5.5) for which $B(\varepsilon, V)$ is continuous at $V = \xi(\varepsilon)$.

It remains to prove that inequality (5.24) holds. By construction we have

$$\frac{\partial B}{\partial V}(\varepsilon, \eta(\varepsilon)) = \frac{1}{n}. \tag{5.26}$$

The differential equation for B in *(B2)* gives us the following, valid for $V \in (\xi(\varepsilon), \eta(\varepsilon))$:

$$\frac{\sigma^2}{2}V^2\frac{\partial^2 B}{\partial V^2}(\varepsilon, V) + (r-\delta)V\frac{\partial B}{\partial V}(\varepsilon, V) - rB(\varepsilon, V) + \rho = 0.$$

Property *(P4)* states that $V \mapsto B(\varepsilon, V)$ is C^1 at $V = \eta(\varepsilon)$. However, it may not be C^2 at $V = \eta(\varepsilon)$. Therefore, taking limits as V increases to $\eta(\varepsilon)$ yields

$$\begin{aligned} \lim_{V\uparrow\eta(\varepsilon)} \left\{ \frac{\sigma^2}{2} V^2 \frac{\partial^2 B}{\partial V^2}(\varepsilon, V) \right\} + (r-\delta)\frac{\eta(\varepsilon)}{n} - r\frac{\eta(\varepsilon)}{n} + \rho &= 0 \\ \Rightarrow \qquad \lim_{V\uparrow\eta(\varepsilon)} \left\{ \frac{\sigma^2}{2} V^2 \frac{\partial^2 B}{\partial V^2}(\varepsilon, V) \right\} &= \delta\frac{\eta(\varepsilon)}{n} - \rho \qquad (5.27) \end{aligned}$$

The bound on $\eta(\varepsilon)$ in (5.5) shows that the right-hand-side of (5.27) is positive. Therefore the function $V \mapsto B(\varepsilon, V)$ is convex for V close to $\eta(\varepsilon)$, and so $B(\varepsilon, V)$ exceeds V/n when V is close to (but less than) $\eta(\varepsilon)$. We also know, from (5.2) and ($B1$), that $B(\varepsilon, V)$ exceeds V/n at $V = \xi(\varepsilon)$. Therefore if $B(\varepsilon, V)$ were less than V/n for some V in the interval $(\xi(\varepsilon), \eta(\varepsilon))$, we would require that it crossed the line V/n (at least) twice in this interval. For this to happen B would have to change from being convex to concave, and then to convex again as V increased from $\xi(\varepsilon)$ to $\eta(\varepsilon)$. This contradicts lemma 5.1 so we deduce that (5.24) holds. □

Lemma 5.2 does not claim the uniqueness of $\xi(\varepsilon)$ and $\eta(\varepsilon)$. Numerical evidence indicates that $\xi(\varepsilon)$ and $\eta(\varepsilon)$ are unique. If, however, there are several pairs $\xi(\varepsilon)$ and $\eta(\varepsilon)$ that satisfy lemma 5.2, we can choose any pair and build the Nash equilibrium from these.

Now we aim to define m^* and the functions ξ and η on the interval $(\varepsilon, m^*]$. It will be easier to work with the functions S and Y rather than with S and B, where Y is defined by

$$Y(m, V) \quad = \quad S(m, V) - B(m, V).$$

As figures 5-2 and 5-3 show, there are two different forms of the solution at $m = \varepsilon$ depending on whether $Y(\varepsilon, V)$ exceeds 0 or not on the interval $V \in (\xi(\varepsilon), \eta(\varepsilon))$. If it does, then we define $\eta(m)$ for $m \in (\varepsilon, m^*]$ to be the root of

$$Y(\varepsilon, V) \quad = \quad 0$$

on the interval $(\xi(\varepsilon), \eta(\varepsilon))$. (As $Y(m, V)$ is negative at $V = \xi(\varepsilon)$ and zero at $V = \eta(\varepsilon)$, the form of $Y(m, V)$ deduced from (5.14) and (5.23) together with lemma 5.1 imply that there is at most one root on this interval.) Otherwise we define $\eta(m) = \eta(\varepsilon)$ for $m \in (\varepsilon, m^*]$. This definition is consistent with *(P5)*, *(P10)*, and conditions (i), (iii) and (iv) of Theorem 5.1.

We can find a differential equation for $Y(m,V)$ on the interval $V \in (\xi(m), \eta(m))$ from *(S2)* and *(B2)*:

$$\mathcal{L}Y + \frac{\delta V - (n - m\tau)\rho}{n - m} = 0 \tag{5.28}$$

The general solution to this differential equation is

$$Y(m,V) = \frac{V - (n - m\tau)\rho/r}{n - m} + a''(m)V^{-\alpha} + b''(m)V^{\beta} \tag{5.29}$$

for any functions $a''(m)$ and $b''(m)$. We can infer from $(B4)$ that

$$Y(m, \eta(m)) = 0$$

and from *(S1)* and *(B1)* that

$$Y(m, \xi(m)) = -\frac{p\xi(m)}{m}.$$

As S and B must both be continuous, Y must be continuous. Therefore we can solve for the functions $a''(m)$ and $b''(m)$:

$$\begin{aligned} a''(m) &= \frac{\eta(m)^{\beta}g(m) + \xi(m)^{\beta}h(m)}{(\xi(m)^{-\alpha}\eta(m)^{\beta} - \xi(m)^{\beta}\eta(m)^{-\alpha})(n-m)} \\ b''(m) &= -\frac{\eta(m)^{-\alpha}g(m) + \xi(m)^{-\alpha}h(m)}{(\xi(m)^{-\alpha}\eta(m)^{\beta} - \xi(m)^{\beta}\eta(m)^{-\alpha})(n-m)}. \end{aligned}$$

where

$$\begin{aligned} g(m) &= (n - m\tau)\rho/r - (np + m(1-p))\xi(m)/m, \\ h(m) &= \eta(m) - (n - m\tau)\rho/r. \end{aligned}$$

For $m \in (\varepsilon, m^*]$ we will define $\xi(m)$ so that $V \mapsto S(m,V)$ is continuous at $\eta(m)$. We define m^* so that *(P6)* and *(P8)* hold. We do this in lemma 5.3.

Lemma 5.3. *There exist m^* and $\xi(m)$ for $\varepsilon < m \le m^*$ satisfying inequalities (5.3), (5.4) and and conditions (i) and (ii) of Theorem 5.1 such that the function $S(m,V)$ is continuous and properties* (P6) *and* (P8) *hold.*

Proof of lemma 5.3. First we will define the default boundary $\xi(m)$ so that $S(m,V)$ is continuous. We have already determined the value of $S(m,\eta(m))$ through (5.15). Equation (5.14) gives the form of $S(m,V)$ for $V \in (\xi(m), \eta(m))$. In lemma 5.2 we proved that there is a unique choice of $\xi(\varepsilon)$ that made $V \mapsto S(\varepsilon, V)$ continuous at $\eta(\varepsilon)$. The same argument can be applied here and so we do not repeat the details. Briefly, if we choose a small enough $\xi(m)$ then $V \mapsto S(m, V)$ will have a negative jump at $\eta(m)$. If we choose a large enough $\xi(m)$ then we have a positive jump at $\eta(m)$. Thus there exists a value of $\xi(m)$ which makes S continuous at $V = \eta(m)$. Uniqueness follows exactly as in lemma 5.2.

We now aim to define m^*. As $Y(\varepsilon, V)$ is negative on $V \in (\xi(\varepsilon), \eta(\varepsilon+))$ we deduce that

$$\frac{\partial Y}{\partial V}(\varepsilon, \eta(\varepsilon+)) \quad \geq \quad 0.$$

If we define m^* by

$$m^* \quad = \quad \inf\{m : \text{the left-derivative of } V \mapsto Y(m, V) \text{ at } \eta(m) \text{ is non-positive}\} \tag{5.30}$$

then lemma 5.1 implies that

$$Y(m, V) \quad < \quad 0$$

for $\varepsilon \leq m \leq m^*$ and $\xi(m) < V < \eta(m)$. Property *(P8)* is therefore satisfied. We have constructed the function $S(m, V)$ to be positive at $V = \eta(m)$. Therefore lemma 5.1 implies that $V \mapsto S(m, V)$ must be convex at $\xi(m)$. We will use this to prove that the bound on $\xi(m)$ in (5.3) is satisfied.

The differential equation for S, valid for $V \in (\xi(m), \eta(m))$, is

$$\frac{\sigma^2}{2} V^2 \frac{\partial^2 S}{\partial V^2}(m, V) + (r - \delta) V \frac{\partial S}{\partial V}(m, V) - rS(m, V) + \frac{\delta V - m\rho'}{n - m} \quad = \quad 0.$$

The function $V \mapsto S(m, V)$ need not be C^2 at $\xi(m)$. However, we know that it is C^1 and that both the function and its derivative are zero at $\xi(m)$. Taking limits as V decreases to $\xi(m)$ yields

$$\lim_{V \downarrow \xi(m)} \left\{ \frac{\sigma^2}{2} V^2 \frac{\partial^2 S}{\partial V^2}(m, V) \right\} + \frac{\delta \xi(m) - m\rho'}{n - m} \quad = \quad 0. \tag{5.31}$$

As $V \mapsto S(m,V)$ is convex at $\xi(m)$ we deduce that

$$\lim_{V \downarrow \xi(m)} \frac{\partial^2 S}{\partial V^2}(m,V) \geq 0$$

which completes the proof. $\square$

It is not obvious that the m^* defined in (5.30) exists. Numerical evidence indicates that for realistic parameter values m^* is close to ε and small with respect to n. We shall proceed with this in mind. Note, however, that if the left-derivative of $V \mapsto Y(m,V)$ at $\eta(m)$ remains positive as m increases to n, lemma 5.3 provides a method for constructing a Nash equilibrium for all m. From now on we will assume that $m^* < n$.

It remains to find the solution when $m > m^*$. We continue to work with S and Y rather than S and B. For V in the interval $(\xi(m), \eta(m))$ the function $Y(m,V)$ is given by the expression in (5.29). We can determine the functions $a''(m)$ and $b''(m)$ in (5.29) in terms of $\eta(m)$ in order to satisfy *(P6)*:

$$\begin{aligned} Y(m,V) &= \frac{V-(n-m\tau)\rho/r}{n-m} + \frac{\beta(n-m\tau)\rho/r-(\beta-1)\eta(m)}{(\alpha+\beta)(n-m)}\left(\frac{V}{\eta(m)}\right)^{-\alpha} \\ &\quad + \frac{\alpha(n-m\tau)\rho/r-(\alpha+1)\eta(m)}{(\alpha+\beta)(n-m)}\left(\frac{V}{\eta(m)}\right)^{\beta} \end{aligned} \tag{5.32}$$

for $V \in (\xi(m), \eta(m))$. Observe that, by construction, this form of $Y(m,V)$ is also valid when $m = m^*$. Lemma 5.4 shows that a solution exists.

Lemma 5.4. *There exist $\xi(m)$ and $\eta(m)$ for $m \in (m^*, n)$ satisfying conditions (i) and (ii) of Theorem 5.1 such that $S(m,V)$ and $Y(m,V)$ are continuous, and* (P8) *holds.*

Figure 5-4 shows how the functions $V \mapsto S(m,V)$ and $V \mapsto Y(m,V)$ typically behave for fixed $m > m^*$. (In figure 5-4 we have $Y(m,V)$ equal to zero for $V \geq \eta(m)$; this will not necessarily be the case.)

Proof of lemma 5.4. It turns out that the boundaries $\xi(m)$ and $\eta(m)$ will be given in terms of the solution to an ordinary differential equation.

In order for $V \mapsto Y(m,V)$ to be continuous at $\xi(m)$ we must have that $Y(m,V)$, as given by the expression in (5.32), tends to $-p\xi(m)/m$ as V tends down to $\xi(m)$. This

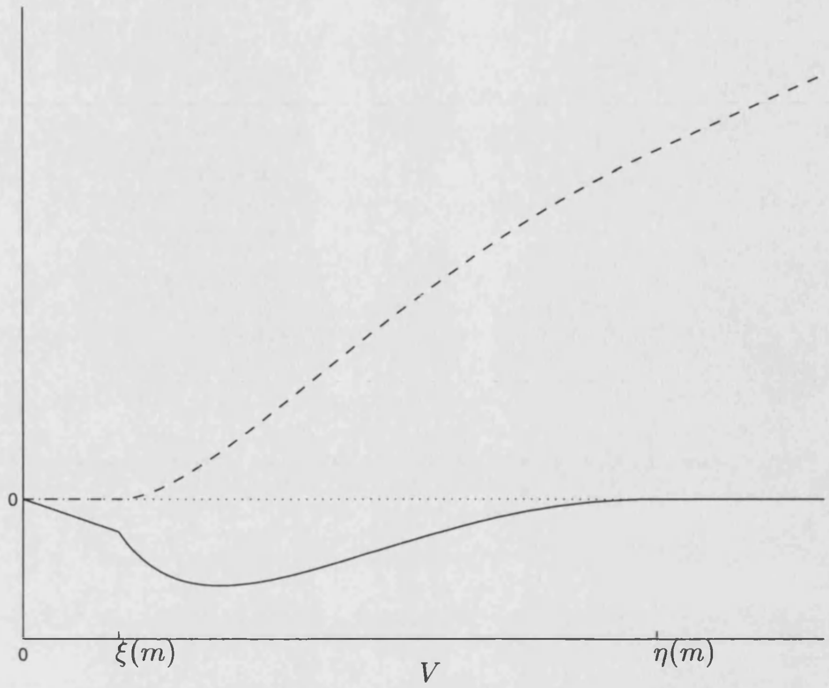

Figure 5-4: Typical behaviour of the functions $S(m, V)$ (dashed line) and $Y(m, V)$ (solid line) for $m \geq m^*$.

gives the following relationship between $\xi(m)$ and $\eta(m)$:

$$
\begin{aligned}
-\frac{p\xi(m)}{m} \;=\; & \frac{\xi(m) - (n - m\tau)\rho/r}{n - m} + \frac{\beta(n - m\tau)\rho/r - (\beta - 1)\eta(m)}{(\alpha + \beta)(n - m)} \left(\frac{\xi(m)}{\eta(m)}\right)^{-\alpha} \\
& + \frac{\alpha(n - m\tau)\rho/r - (\alpha + 1)\eta(m)}{(\alpha + \beta)(n - m)} \left(\frac{\xi(m)}{\eta(m)}\right)^{\beta}. \qquad (5.33)
\end{aligned}
$$

We are unable to write $\xi(m)$ explicitly in terms of $\eta(m)$, or vice versa, from this equation. In fact, it is not immediately clear that $\xi(m)$ uniquely determines $\eta(m)$, or vice versa. We will shortly avoid this problem by changing variables. For now, however, notice that *(P11)* implies that

$$\frac{\partial S}{\partial m}(m, \eta(m)) \;=\; 0. \qquad (5.34)$$

As $S(m, V)$ is continuous we could apply this differential equation to the expression for $S(m, V)$ in (5.14). This would give an expression for the derivative of $\xi(m)$ in terms of $\xi(m)$ and $\eta(m)$. Given the initial value $\xi(m^*)$ we would be able to find $\xi(m)$ and $\eta(m)$ for $m \in (m^*, n)$ through this differential equation and (5.33). This is essentially the method we use. However, we change variables to obtain an explicit differential equation and so avoid the difficulties mentioned above.

We define $\theta(m)$ by

$$\theta(m) \;=\; \frac{\xi(m)}{\eta(m)} \tag{5.35}$$

and work with $\theta(m)$ and $\eta(m)$ rather than with $\xi(m)$ and $\eta(m)$. The solution at m^* gives us the value of $\theta(m^*)$ and we will find that we can obtain an explicit differential equation for $\theta(m)$. Replacing $\xi(m)$ with $\theta(m)\eta(m)$ in (5.33) yields

$$\begin{aligned} -\frac{p\theta(m)\eta(m)}{m} \;&=\; \frac{\theta(m)\eta(m)-(n-m\tau)\rho/r}{n-m}+\frac{\beta(n-m\tau)\rho/r-(\beta-1)\eta(m)}{(\alpha+\beta)(n-m)}\theta(m)^{-\alpha} \\ &+\frac{\alpha(n-m\tau)\rho/r-(\alpha+1)\eta(m)}{(\alpha+\beta)(n-m)}\theta(m)^{\beta}. \end{aligned}$$

We can now write $\eta(m)$ in terms of $\theta(m)$:

$$\eta(m) \;=\; \frac{(n-m\tau)\left(\beta\theta(m)^{-\alpha}+\alpha\theta(m)^{\beta}-\alpha-\beta\right)\rho/r}{(\beta-1)\theta(m)^{-\alpha}+(\alpha+1)\theta(m)^{\beta}-(\alpha+\beta)\theta(m)(np+m(1-p))/m} \tag{5.36}$$

We also rewrite (5.14) in terms of $\eta(m)$ and $\theta(m)$:

$$\begin{aligned} S(m,V) \;&=\; \frac{V-m\rho'/r}{n-m}+\frac{\beta m\rho'/r-(\beta-1)\theta(m)\eta(m)}{(\alpha+\beta)(n-m)}\left(\frac{V}{\theta(m)\eta(m)}\right)^{-\alpha} \\ &+\frac{\alpha m\rho'/r-(\alpha+1)\theta(m)\eta(m)}{(\alpha+\beta)(n-m)}\left(\frac{V}{\theta(m)\eta(m)}\right)^{\beta} \end{aligned} \tag{5.37}$$

Equation (5.36) allows us to eliminate $\eta(m)$ from (5.37), giving an expression for $S(m,V)$ in terms of $\theta(m)$ only. Applying the differential equation (5.34) to this form of $S(m,V)$ then gives an explicit differential equation for $\theta(m)$. We have an initial value $\theta(m^*)$ and so we can find $\theta(m)$ for $m \in (m^*,n)$. For each m, equations (5.36) and then (5.35) give $\eta(m)$ and then $\xi(m)$ from $\theta(m)$.

Next we prove that *(P8)* holds. We are also able to find a lower bound on $\eta(m)$. From (5.28) we have the following differential equation for Y:

$$\frac{\sigma^2}{2}V^2\frac{\partial^2 Y}{\partial V^2}(m,V)+(r-\delta)V\frac{\partial Y}{\partial V}(m,V)-rY(m,V)+\frac{\delta V-(n-m\tau)\rho}{n-m} \;=\; 0.$$

We have defined $V \mapsto Y(m,V)$ so that both the function and its left derivative are zero at $\eta(m)$. Therefore taking limits as V increases to $\eta(m)$ gives

$$\lim_{V\uparrow\eta(m)}\left\{\frac{\sigma^2}{2}V^2\frac{\partial^2 Y}{\partial V^2}(m,V)\right\}+\frac{\delta\eta(m)-(n-m\tau)\rho}{n-m} \;=\; 0.$$

Now $Y(m,\xi(m))$ is negative and so lemma 5.1 implies that we must have

$$Y(m,V) \;<\; 0$$

for $V \in (\xi(m),\eta(m))$. This is equivalent to *(P8)*. We also deduce that $Y(m,V)$ must be concave in V at $V=\eta(m)$. Therefore

$$\lim_{V\uparrow\eta(m)}\frac{\partial^2 Y}{\partial V^2}(m,V) \;\leq\; 0$$

and the bound

$$\eta(m) \;\geq\; (n-m\tau)\frac{\rho}{\delta} \tag{5.38}$$

follows $\square$

We have proved that the boundaries $\xi(m)$ and $\eta(m)$ and the functions S and B satisfy most of the premises of Theorems 5.1 and 5.2. However, lemma 5.4 does not claim that the bound on $\xi(m)$ in (5.3) holds, that we can take $\bar{m}=n$ in *(P7)*, or that condition (*iv*) of Theorem 5.1 holds. We now prove the first two of these properties.

Lemma 5.5. *With the functions ξ and S defined through lemma 5.4, inequality (5.3) is satisfied and we may take $\bar{m}=n$ in* (P7).

Proof of lemma 5.5. The form of S in (5.9) together with lemma 5.1 shows that on the interval $(\xi(m),\eta(m))$, the function $V\mapsto S(m,V)$ can have at most two turning points and cannot have a maximum and then a minimum. We defined the function S so that $\frac{\partial S}{\partial V}(m,\xi(m))=0$. Therefore if

$$S(m,\eta(m)) \;>\; 0 \tag{5.39}$$

we deduce that $S(m,V)$ is positive for all $V\in(\xi(m),\eta(m))$. This would imply that $V\mapsto S(m,V)$ is convex at $\xi(m)$. We could then deduce from (5.31) that the bound on $\xi(m)$ in (5.3) holds. It is therefore sufficient to prove that (5.39) holds for all $m\in(m^*,n)$.

By construction (5.39) holds for $m \leq m^*$. Suppose that $\hat{m}$ defined by

$$\hat{m} \quad = \quad \inf \{m : S(m, \eta(m)) \leq 0\}$$

is less than n. The continuity of $S(m, V)$ and $\eta(m)$ imply that $\hat{m} > m^*$ and that

$$S(\hat{m}, \eta(\hat{m})) \quad = \quad 0. \tag{5.40}$$

and that *(P7)* holds if we take $\bar{m} = \hat{m}$. Theorem 5.2 therefore shows that $S(m, V)$ corresponds to the value of the share for $m < \hat{m}$. Now consider the net present value of the cashflows from a share if $m_0 = \hat{m}$ and $V_0 = \eta(\hat{m})$. The bound on $\eta(m)$ in (5.38) shows that the dividend rate to a shareholder, given by

$$\frac{\delta V - m\rho'}{n - m},$$

is positive at time zero. Under Π_B a bondholder will convert immediately and so $m_{0+} < m_0 = \hat{m}$ and hence $S(m_t, V_t)$ will be non-negative at $t = 0+$. The function $S(m_{0+}, V_{0+})$ corresponds to the share price at (m_{0+}, V_{0+}). Therefore we must have that the net time-0 value of the cashflow to a shareholder is strictly positive, contradicting (5.40). This proves that (5.39) holds for all $m \in (m^*, n)$. □

The final property that we need is condition (iv) of Theorem 5.1, for $m \geq m^*$. This property is required as S and B are defined through (5.6) and (5.7) when $\eta(m) < V \leq \eta(m^*)$ and they may not be continuous in m across the boundary $\eta(m)$ if $\eta(m)$ is not decreasing. For m on the interval (m^*, n) the boundary $\eta(m)$ is defined in terms of $\theta(m)$ by (5.36). The function $\theta(m)$ is the solution to a differential equation, and it is difficult to obtain properties of $\theta(m)$. We leave condition (iv) of Theorem 5.1 as an assumption. It is satisfied in all numerical examples we have looked at.

5.4 Smooth pasting

Property *(S3)* of definition 5.1 states that the function $V \mapsto S(m, V)$ is C^1 at the default boundary. Property *(P4)* of Theorem 5.2 requires that $V \mapsto B(\varepsilon, V)$ is C^1 at the conversion boundary, and property *(P6)* of the same theorem states that the left derivative of $V \mapsto Y(m, V)$ is equal to zero at the conversion boundary. All these conditions are examples of smooth pasting conditions. In section 5.3 we have proved that we have a Nash equilibrium. In this section we attempt to justify the smooth pasting conditions more intuitively and explain why a smooth pasting condition corresponds to the optimal choice of a boundary.

For this purpose we consider the default boundary $\xi(m)$. The location of this boundary is determined by the firm and should therefore be chosen so as to maximise the share value. At and below the default boundary the share value will be zero. Above the default boundary there will be an interval in which all agents take no action. In this interval the share price solves the differential equation in *(S2)* of definition 5.1 and so is given by (5.9). As can be seen from (5.9), this differential equation does not determine S explicitly — there are two unknowns $a(m)$ and $b(m)$ to be chosen. Typically we have a boundary condition, independent of the choice of $\xi(m)$, which gives one constraint on the choice of these functions. For example, Leland (1994) studies the pricing of straight debt (that is, bonds issued without the convertible feature) and obtains a boundary condition from the behaviour of the price of this debt as V tends to infinity. In our context the boundary condition is given by the price of the share at the conversion boundary $\eta(m)$. The boundary $\eta(m)$ is determined by the bondholders and the share price at this boundary is independent of the choice of $\xi(m)$. Note that we do not require that the choice of conversion boundary $\eta(m)$ is optimal for the bondholders, only that the choice of conversion boundary fixes the share price there.

The choice of $\xi(m)$ gives a second boundary condition, as $S(m, V)$ must be continuous at $V = \xi(m)$, and so determines the functions $a(m)$ and $b(m)$. Suppose, for some choice of $\xi(m)$, that the share price between $\xi(m)$ and $\eta(m)$ is given by

$$S(m, V) \quad = \quad f(V)$$

and that this function $f(V)$ does not smooth paste to zero at $V = \xi(m)$. Then consider the function

$$\tilde{f}(V) \quad = \quad f(V) + x\left\{\left(\frac{V}{\eta(m)}\right)^{-\alpha} - \left(\frac{V}{\eta(m)}\right)^{\beta}\right\}$$

for some positive x chosen small enough. Note that the expression in the curly brackets is positive for $V < \eta(m)$ and zero at $V = \eta(m)$. The function $\tilde{f}(V)$ satisfies the differential equation in *(S2)* and the boundary condition at $V = \eta(m)$ and so is a valid choice for the share price function, provided that it reaches zero at some value of V. We would then choose the default boundary $\tilde{\xi}(m)$ to be this value of V.

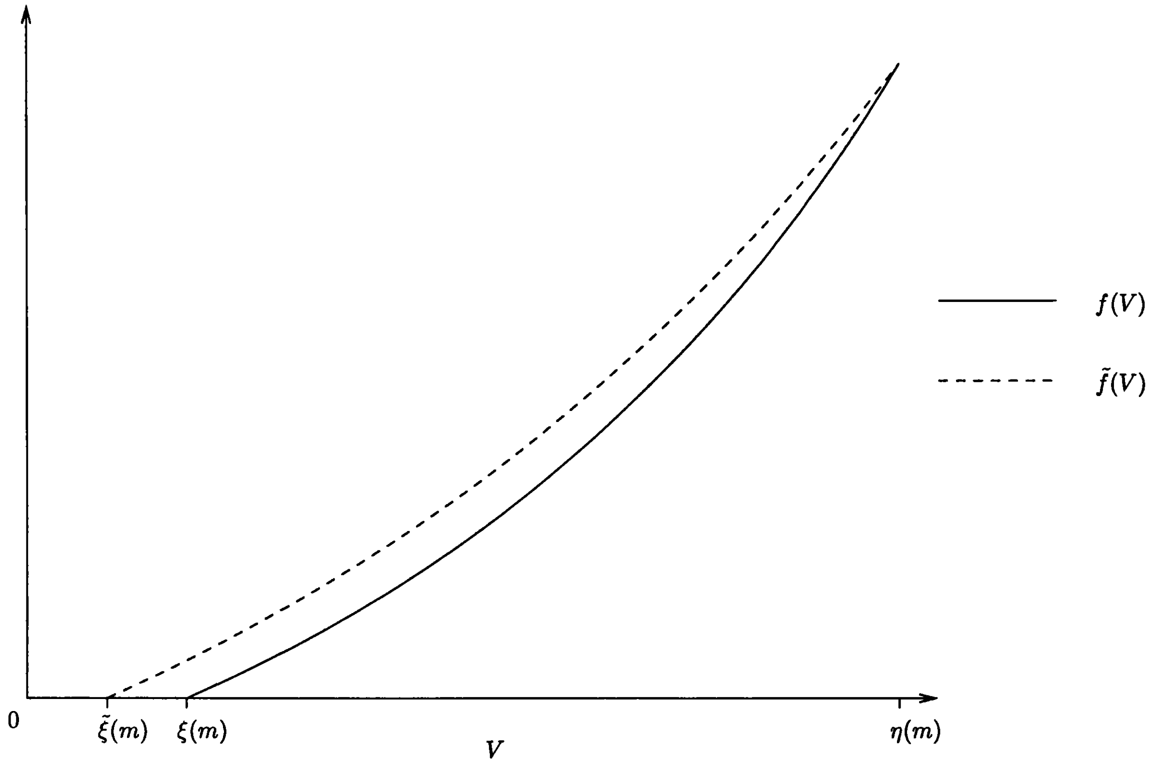

Figure 5-5: The smooth pasting condition

As $f(V)$ does not smooth paste to zero it must cross zero at $\xi(m)$, and be negative either just to the left or just to the right of $\xi(m)$. Therefore if we choose x small enough, $\tilde{f}(V)$ will reach zero. Furthermore, we have

$$\tilde{f}(V) \quad > \quad f(V)$$

for $V < \eta(m)$, and so the new choice of default boundary increases the share price on this range. Figure 5-5 illustrates this for a function $f(V)$ with positive gradient at $\xi(m)$.

We conclude that if $S(m, V)$ does not smooth paste to zero at $V = \xi(m)$ it corresponds to a suboptimal choice of $\xi(m)$. If $S(m, V)$ does smooth paste to zero at $V = \xi(m)$ then any different choice of $\xi(m)$ will decrease the share price. Similar arguments can be used to justify the other smooth pasting conditions we have used.

5.5 Parameter reduction

There is a certain amount of redundancy in the parameters that we have used. Consider (5.1), which gives the dynamics of the firm value process V_t. As

$$\sqrt{\lambda}\sigma W_t \quad =_d \quad \sigma W_{\lambda t}$$

for $\lambda > 0$ (where $=_d$ denotes equality in distribution), we see that multiplying the volatility σ of the firm value process by $\sqrt{\lambda}$ and multiplying r and δ by λ can be thought of as speeding up the firm value process by a factor of λ. Suppose also that we multiply the rates at which the firm pays out coupons by the same factor. The net effect of these changes would be the same as that of scaling time. As our model does not depend explicitly on time, we would not have changed the solution. Thus the default and conversion boundaries would remain unchanged. Changing notation slightly to express the dependence of the share price and bond price on some of the other parameters in the model, we obtain the following relationships:

$$\begin{aligned} S(\rho, r, \delta, \sigma; m, V) &= S(\lambda\rho, \lambda r, \lambda\delta, \sqrt{\lambda}\sigma; m, V) \\ B(\rho, r, \delta, \sigma; m, V) &= B(\lambda\rho, \lambda r, \lambda\delta, \sqrt{\lambda}\sigma; m, V). \end{aligned}$$

Secondly, suppose that the process V_t describes the firm value for each of μ identical firms, each as described in the model, and each with the same number of outstanding convertible bonds. The pattern of behaviour in each of these firms will be the same — if bonds are converted in one firm, the same will happen in each of the others. Now suppose that the μ firms are considered as one. This combined firm fits into the model that we have used. The total number of shares and bonds in the combined firm is μn, the state of the system can be described by $(\mu m, \mu V)$ and the final $\mu\varepsilon$ bonds must be converted together. The holder of a bond or share in one of the μ firms can be thought of as the holder of a bond or a share respectively in the combined firm, and the price of the assets will not have changed. Therefore, again using different notation to emphasise the dependence of the prices on the parameters of interest, we deduce that

$$\begin{aligned} S(n, \varepsilon; m, V) &= S(\mu n, \mu\varepsilon; \mu m, \mu V) \\ B(n, \varepsilon; m, V) &= B(\mu n, \mu\varepsilon; \mu m, \mu V) \end{aligned}$$

The default and conversion boundaries will be scaled by μ, as we would expect.

Finally, suppose that the coupon rate and firm value process are both multiplied by ν. This also has the effect of increasing the dividend rate by a factor of ν. This transformation is equivalent to changing the units of currency being used (at a fixed rate of exchange) and so has no fundamental effect on the solution. The boundaries, measured in the new currency, will change by a factor of ν and the share and bond prices will both increase by the same factor. Thus

$$\begin{aligned} S(\rho; m, V) &= \frac{1}{\nu} S(\nu\rho; m, \nu V) \\ B(\rho; m, V) &= \frac{1}{\nu} B(\nu\rho; m, \nu V). \end{aligned}$$

Combining these three identities yields

$$\begin{aligned} S(n, \varepsilon, \rho, r, \delta, \sigma; m, V) &= \frac{\rho}{\sigma^2} S\left(1, \frac{\varepsilon}{n}, 1, \frac{r}{\sigma^2}, \frac{\delta}{\sigma^2}, 1; \frac{m}{n}, \frac{\sigma^2 V}{n\rho}\right) \\ B(n, \varepsilon, \rho, r, \delta, \sigma; m, V) &= \frac{\rho}{\sigma^2} B\left(1, \frac{\varepsilon}{n}, 1, \frac{r}{\sigma^2}, \frac{\delta}{\sigma^2}, 1; \frac{m}{n}, \frac{\sigma^2 V}{n\rho}\right). \end{aligned}$$

The solution that we have consists of share and bond price functions, together with policies concerning default and conversion. From the specification of the model, it is clear that ε/n is a fundamental parameter of the model. We have found that the share price and bond price depend upon the parameters α and β, which are themselves defined in terms of the solutions to the quadratic in (5.11). As r and δ appear independently in this quadratic, the solution fundamentally depends on both of these parameters. Therefore no further parameter reduction is possible.

5.6 Numerical examples

In this section we choose two sets of parameter values and compute the solution in each case. We take $n = 1$ in both examples — as discussed in section 5.5 this does not result in any loss of generality.

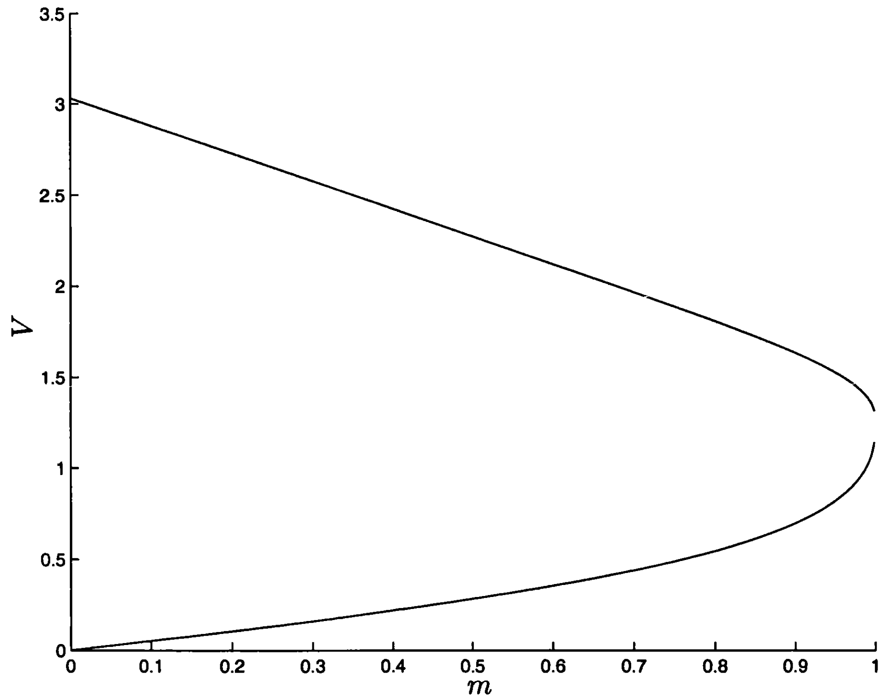

Figure 5-6: The conversion boundary (upper curve) and default boundary (lower curve).

Example 1

Figure 5-6 shows the default and conversion boundaries for the following parameter values:

$$n = 1, \quad \varepsilon = 1/800, \quad \rho = 5\%, \quad r = 4\%,$$
$$\delta = 2\%, \quad \sigma = 10\%, \quad p = 0.2, \quad \tau = 0.5.$$

From these parameter values we calculate that

$$\alpha = 4.70, \quad \beta = 1.70.$$

In this example the functions $S(\varepsilon, V)$ and $B(\varepsilon, V)$ look like those in figure 5-2 rather than figure 5-3, that is the share price exceeds the bond price close to $\eta(\varepsilon)$. We find that

$$\begin{aligned} \varepsilon &= 0.001250, & \eta(\varepsilon) &= 3.031738, \\ m^* &= 0.002499, & \eta(m^*) &= 3.027951. \end{aligned}$$

We see that m^* is roughly twice ε, and the jump in the conversion boundary is less than 0.2%.

In order to see how significant the choice of ε is, I repeated the calculation four further times. All parameters except ε were unchanged, and ε took the values 1/400, 1/200, 1/100, and 1/50. In each case m^* was roughly double ε. The curves of the boundaries that these calculations produce are so similar that a graph is not useful. For example, the largest difference between any two corresponding values of $\xi(m)$ or $\eta(m)$ was less than 0.01%.

Figures 5-7 and 5-8 show how the share price and the bond price vary with V for five different values of m. For low values of V the share price is decreasing in m, and for large values of V it is increasing in m. This is perhaps not surprising as the dividend rate exhibits the same behaviour:

$$\frac{\partial}{\partial m}\left(\frac{\delta V - m\rho'}{n-m}\right) = \frac{\delta V - n\rho'}{(n-m)^2}.$$

The value of a perpetual risk-free bond paying coupon ρ is ρ/r which is 1.25 in this example. For large values of V there is a high probability that a bond will be converted and the probability of default is low. The bond price is therefore close to that of the share, as the graphs show. As V decreases the curves seem to flatten with the bond value close to 1.25, particularly when m is low. Then as V decreases further, the price of the bond approaches its value at default.

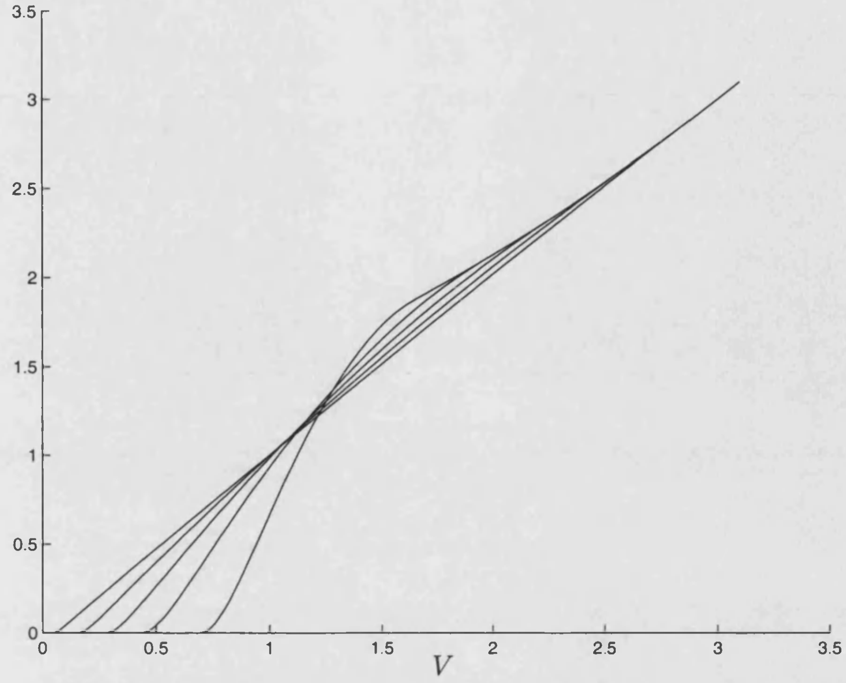

Figure 5-7: Share price as a function of firm value. The five curves correspond to $m = 0.1$, $m = 0.3$, $m = 0.5$, $m = 0.7$ and $m = 0.9$.

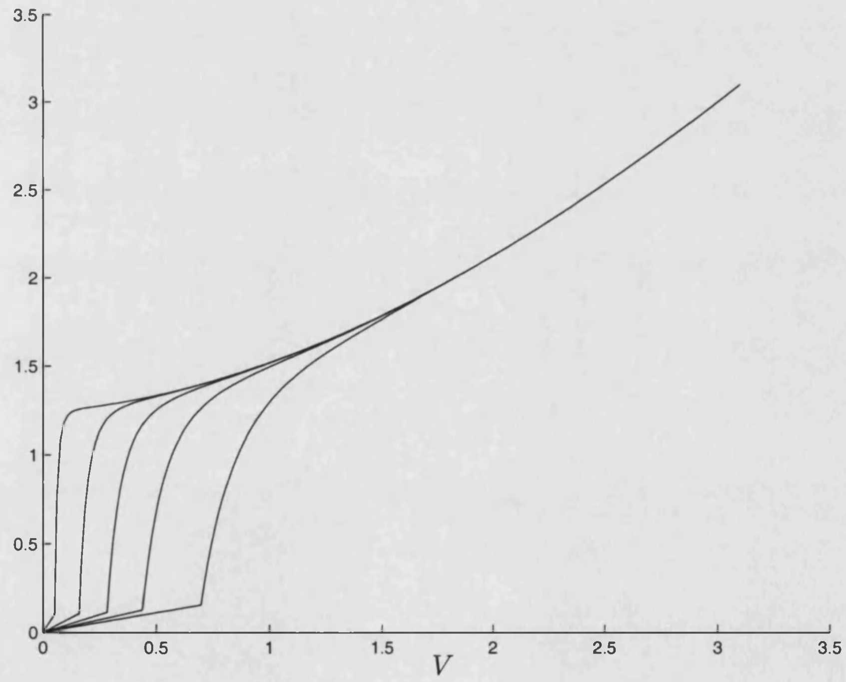

Figure 5-8: Bond price as a function of firm value. The five curves correspond to $m = 0.1$, $m = 0.3$, $m = 0.5$, $m = 0.7$ and $m = 0.9$.

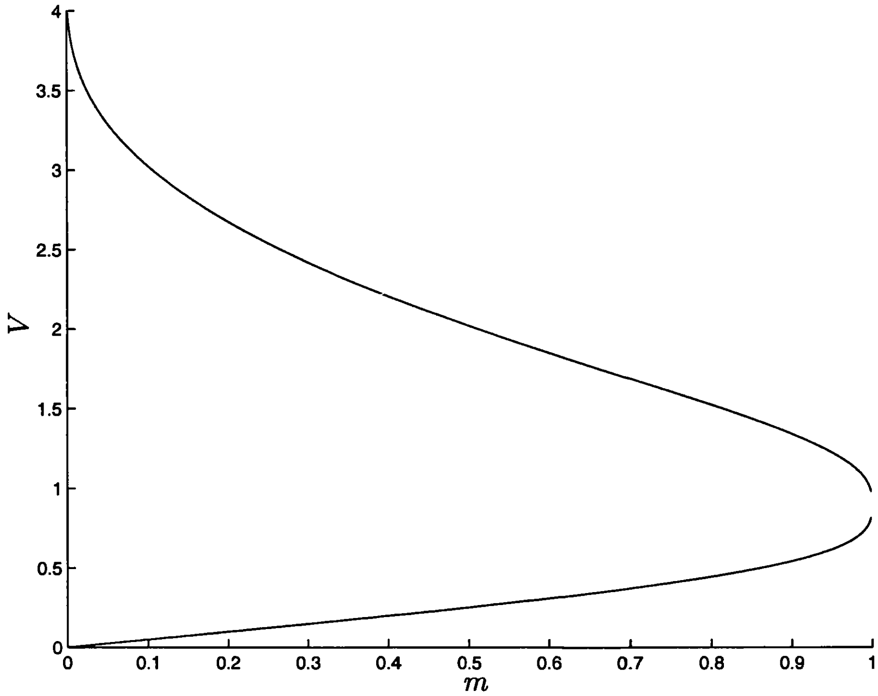

Figure 5-9: The conversion boundary (upper curve) and default boundary (lower curve).

Example 2

The parameters for this example are

$$n = 1, \quad \varepsilon = 1/800, \quad \rho = 7\%, \quad r = 3\%,$$
$$\delta = 7\%, \quad \sigma = 30\%, \quad p = 0.5, \quad \tau = 0.1.$$

From these parameter values we calculate that

$$\alpha = 0.304, \quad \beta = 2.19.$$

Figure 5-9 shows the default and conversion boundaries. Again, the functions $S(\varepsilon, V)$ and $B(\varepsilon, V)$ look like those in figure 5-2 rather than figure 5-3, as the values below indicate.

$$\begin{aligned} \varepsilon &= 0.001250, & \eta(\varepsilon) &= 3.986529, \\ m^* &= 0.001264, & \eta(m^*) &= 3.984980. \end{aligned}$$

Here m^* is only one percent greater than ε. In the first example, $\eta(m)$ appeared to fall away from $\eta(\varepsilon)$ linearly. In this example $\eta(m)$ falls away at a faster rate. As with the first example, altering the value of ε within reasonable bounds does not make a significant difference to the boundaries.

Figures 5-10 and 5-11 show how the share price and bond price vary with V. A difference between this example and example 1 is that for the five values of m plotted, the share price appears not to exceed V/n. A reason for this could be that the tax rebate to the firm is significantly less here. The flattening of the bond prices is less pronounced in this example.

Now we attempt to understand how the default and conversion boundaries behave when m is close to n. When $m = n - x$ the dividend rate is

$$\frac{\delta V - (n-x)\rho'}{x}.$$

For small enough x, this expression will be large in magnitude, with sign depending on whether

$$\delta V - n\rho'$$

is negative or not. With a large negative dividend we would expect the firm to default. With a large positive dividend we would expect the optimal policy for the bondholders to be to convert — the dividend rate will greatly exceed the coupon rate. Therefore we expect that both the default boundary and the conversion boundary will be close to $n\rho'/\delta$, whilst still satisfying the inequalities in (5.3) and (5.38). This is consistent with the behaviour that we see in the numerical examples.

5.7 Discussion

In this section I attempt to explain some of the modelling choices that I have made in this chapter. The reason why it is specified that the last ε bonds are converted together may not be clear. Removing this constraint would lead us to look for a solution using the methods of lemma 5.4 for all m. In particular, we might hope to be able to solve a differential equation for the default and conversion boundaries for all m. The initial condition for this first-order differential equation could be supplied by the solution at

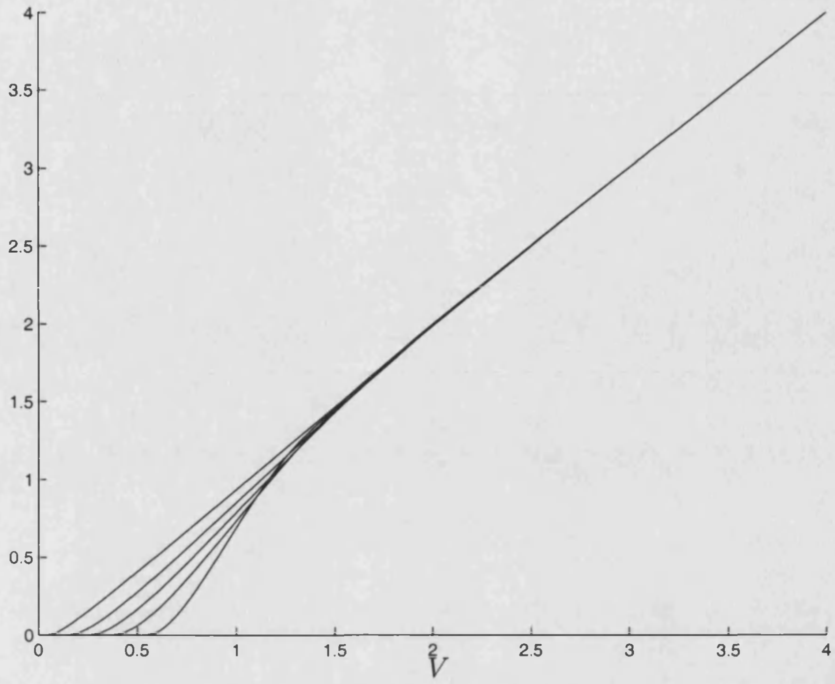

Figure 5-10: Share price as a function of firm value. The five curves correspond to $m = 0.1$, $m = 0.3$, $m = 0.5$, $m = 0.7$ and $m = 0.9$.

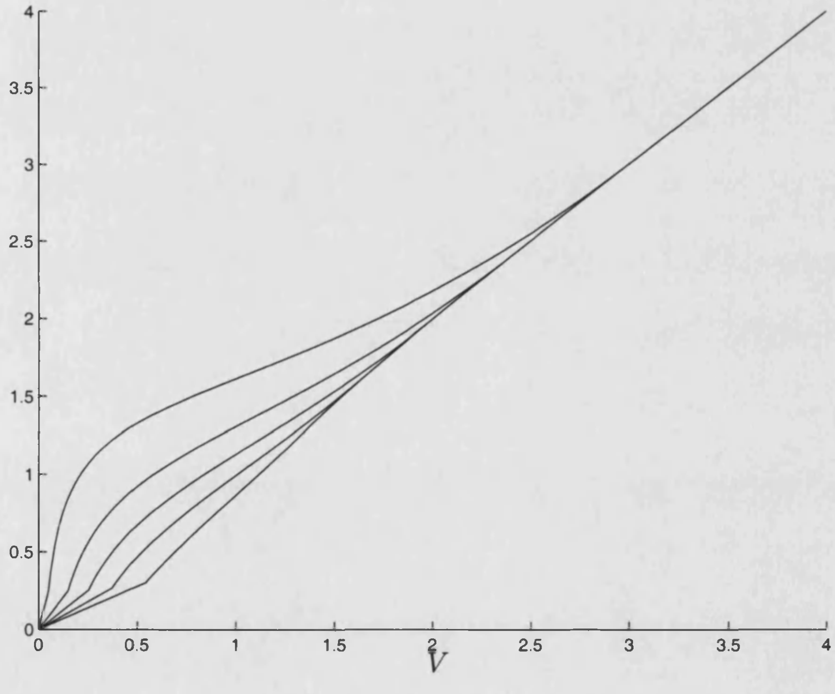

Figure 5-11: Bond price as a function of firm value. The five curves correspond to $m = 0.1$, $m = 0.3$, $m = 0.5$, $m = 0.7$ and $m = 0.9$.

$m = 0$, that is, by $\xi(0) = 0$. The difficulty with this method lies with this initial condition. It turns out that the differential equation has a singularity at $m = 0$ and so the existence and uniqueness of a solution do not follow easily.

Having found these difficulties with the approach with a continuous process m_t, a natural next step would be to completely discretise the problem with respect to m_t. We would specify that bonds and shares could only be owned and converted in integer multiples of ε for some small ε. The process m_t would then be an integer multiple of ε at all times. In appendix C we look at this version of the problem in more detail. Minor changes are needed to the definitions of S and B as the domain of these functions has changed. For example the definition of $S(m, V)$ above the conversion boundary would be

$$S(m,V) \quad = \quad S(m-\varepsilon,V) \quad \text{ for } V > \eta(m).$$

The values of $\xi(\varepsilon)$ and $\eta(\varepsilon)$ can be found exactly as in lemma 5.2. The most significant difference in the solution is that the evolution of the default and conversion boundaries can no longer given in terms of a differential equation. We find instead that the conversion boundary $\eta(m)$ for $m \geq 2\varepsilon$ is a root of the equation

$$Y(m-\varepsilon,V) \quad = \quad 0 \tag{5.41}$$

on the interval $V \in [\xi(m-\varepsilon), \eta(m-\varepsilon)]$. (One root is always given by $V = \eta(m-\varepsilon)$.) This is similar to the way in which we found $\eta(m)$ on the interval $m \in (\varepsilon, m^*]$ from the function $V \mapsto Y(\varepsilon, V)$ in section 5.3. We can therefore build the conversion and default boundaries iteratively.

The condition corresponding to *(P8)* becomes

$$S(m-\varepsilon,V) \quad < \quad B(m,V) \quad \text{ for } V \in (\xi(m), \eta(m)). \tag{5.42}$$

Intuitively this condition states that a bondholder who converts before V reaches the conversion boundary cannot increase the net value of his future cashflows. A difficulty with the completely discretised version of the problem comes from this condition. It is difficult to prove that this inequality holds. One reason for this is that it is hard to obtain information about the roots to (5.41).

It is also numerically difficult to find solutions to (5.41) and to verify (5.42). For some values of m a root to (5.41) exists on the interval $V \in (\xi(m - \varepsilon), \eta(m - \varepsilon))$, for other values the only root is $V = \eta(m - \varepsilon)$. Typically the partial derivative

$$\frac{\partial Y}{\partial V}(m - \varepsilon, V)$$

will be close to zero at $V = \eta(m - \varepsilon)$ and if a root to (5.41) exists on $V \in (\xi(m - \varepsilon), \eta(m-\varepsilon))$ it is close to $\eta(m-\varepsilon)$. Repeating numerical examples with varying degrees of accuracy showed that the root-finding routines used were not stable.

By construction we have equality in (5.42) when $V = \eta(m)$. Verifying (5.42) numerically can be attempted by comparing the left partial derivatives with respect to V of each side at $V = \eta(m)$. Again, the numerical results that I obtained depended on the accuracy to which calculations were performed and were not stable as the accuracy increased. We construct the solution iteratively in m and so any errors will accumulate. This is likely to contribute to the effects just discussed.

An advantage of the completely discretised version of the problem over the version we have used in this chapter is that by construction we have a decreasing conversion boundary. This in part compensates for the difficulties above. The version of the problem presented in this chapter is in some sense a hybrid of the discrete and continuous versions. We use the discrete version of the problem to avoid the difficulties with the initial condition for the differential equation, and then revert to the continuous version of the problem. The solution on the interval $(\varepsilon, m^*]$ effectively joins these two solutions.

Finally, despite the problems described above, the default and conversion boundaries that were numerically calculated in the discrete version of the problem were reasonably stable. Even though the values of m for which a root to (5.41) is found on the interval $V \in (\xi(m - \varepsilon), \eta(m - \varepsilon))$ depend on the accuracy used, the values of the boundaries varies very little. This is partly because any roots to (5.41) are close to $\eta(m-\varepsilon)$. If we compare numerically the boundaries obtained in the discrete version of the problem to those obtained in section 5.6 we find that the difference is very small.

5.8 Conclusions

In this chapter we have analysed a model of a firm with an issue of perpetual convertible debt. We have endogenously found the optimal terms under which the firm defaults, and the optimal conditions for bondholders to convert.

Each bondholder aims to maximise the value of the cash flow he will receive in the future. The cash flow for a bondholder consists of fixed coupon payments whilst the firm is solvent, and a one-off terminal payment should the firm default. The cash flow for a shareholder consists of a dividend stream at a variable rate whilst the firm is solvent, but no payments in the event of default. The dividend rate is positively correlated with the value of the firm. Thus if the value of the firm is high enough, a share is more attractive than a bond. When the firm value is low, the fixed coupons and the payment received in the event of default make a bond more valuable than a share. In deciding whether to convert, a bondholder compares these two future cash flows.

The management of the firm chooses the timing of default in order to maximise value of the future cash flow to the shareholders. The policies of the management of the firm and of the bondholders thus depend on how they expect each other to act. We have found a Nash equilibrium which gives us optimal behaviour for all agents.

One feature of this Nash equilibrium is that the bondholders do not all convert simultaneously. When a bondholder converts, we find that the dividend rate per shareholder decreases. By itself, this would have the effect of decreasing the share price. However, the total coupon payments also decrease and this has the effect of reducing the likelihood of default. The combination of these changes has no net effect on the share price. We also find that the bond price does not change at conversion. However, we often find that further immediate conversion would result in a decrease in share price. This is because the effect of a further decrease in dividend would not be compensated for fully by the decrease in the likelihood of default that would occur. The optimal behaviour in this case is for the remaining bondholders to wait for the firm value to increase before converting.

We have assumed that the bond holding is diffuse and have not allowed any collaboration amongst the bondholders when $m > \varepsilon$. Altering these assumptions would change the nature of the optimal behaviour. For example, consider the situation when the firm value is very close to the default boundary. Here the bond price will be close to pV/m. Should all the bondholders convert, then each would hold a share worth V/n. If m is large enough that

$$p \quad < \quad m/n$$

(and the firm value is close enough to the default boundary) then the bondholders would all benefit from this behaviour. However, our solution shows that any single bondholder would lose out by converting if at least some of the remaining bondholders did not also convert. Considering the problem where the bondholders and/or subsets of the bondholders can collaborate would be complex.

We have modelled perpetual debt rather than finite maturity debt. As a result, we are able to find explicit expressions for the share price and bond price in terms of the default and conversion boundaries. These boundaries can be found numerically. In the case of finite maturity debt, the share price and bond price would be functions of time as well as of the number of remaining convertibles and the firm value. Similarly, the default and conversion boundaries would become functions of time and the number of remaining bonds. It seems unlikely that there would be explicit expressions for the share price and bond price, and so the partial differential equations that arise would have to be solved numerically.

The model we have used could be extended to consider callable convertible debt. A policy for the management of the firm would then additionally specify the conditions under which the firm should call the bonds. For fixed m the solution would no longer necessarily consist of a default boundary and a conversion boundary but may include a call boundary.

Appendix A

Standard barrier option price

The explicit formula for the price of a standard up-and-out barrier option given in Conze & Viswanathan (1991) is reproduced below. We use the symbol Φ to denoted the standard normal distribution function.

$$\begin{aligned}
\Psi(b) = & \left\{ \Phi\left(\frac{b - r - \frac{\sigma^2}{2}}{\sigma}\right) - e^{b+\frac{2rb}{\sigma^2}} \Phi\left(\frac{-b - r - \frac{\sigma^2}{2}}{\sigma}\right) \right. \\
& \left. - \Phi\left(\frac{\log k - r - \frac{\sigma^2}{2}}{\sigma}\right) + e^{b+\frac{2rb}{\sigma^2}} \Phi\left(\frac{-2b + \log k - r - \frac{\sigma^2}{2}}{\sigma}\right) \right\} \\
& - ke^{-r} \left\{ \Phi\left(\frac{b - r + \frac{\sigma^2}{2}}{\sigma}\right) - e^{-b+\frac{2rb}{\sigma^2}} \Phi\left(\frac{-b - r + \frac{\sigma^2}{2}}{\sigma}\right) \right. \\
& \left. - \Phi\left(\frac{\log k - r + \frac{\sigma^2}{2}}{\sigma}\right) + e^{-b+\frac{2rb}{\sigma^2}} \Phi\left(\frac{-2b + \log k - r + \frac{\sigma^2}{2}}{\sigma}\right) \right\}
\end{aligned}$$

Appendix B

Correlation of $g(X, b)$ and τ

In the risk-neutral probability the payoff of a standard barrier option $g(X, b)$ is a function of X_1 and $X_{\tilde{t}}$. We will look at $E\left[\exp\left(\alpha X_1 + \beta X_{\tilde{t}} + \gamma \frac{\tau}{\sqrt{\delta}}\right)\right]$ to assess the correlation between $g(X, b)$ and τ.

$$
\begin{aligned}
& E\left[\exp\left(\alpha X_1 + \beta X_{\tilde{t}} + \gamma \frac{\tau}{\sqrt{\delta}}\right)\right] \\
= \; & E\left[E\left[\exp\left(\alpha X_1 + \beta X_{\tilde{t}} + \gamma \frac{\tau}{\sqrt{\delta}}\right) \middle| \left(X_{j\delta}, \sup_{(j-1)\delta \le u \le j\delta} X_u, j = 1, \ldots, N\right)\right]\right] \\
= \; & E\left[\exp(\alpha X_1 + \beta X_{\tilde{t}}) E\left[\exp\left(\gamma \frac{\tau}{\sqrt{\delta}}\right) \middle| \left(X_{j\delta}, \sup_{(j-1)\delta \le u \le j\delta} X_u, j = 1, \ldots, N\right)\right]\right]
\end{aligned}
$$

Consider the inner expectation in this expression. The Markov property for X_t implies that, of the conditioning variables, τ depends only on $X_{(\tilde{j}-1)\delta}$, $\sup_{(\tilde{j}-1)\delta \le u \le \tilde{j}\delta} X_u = X_{\tilde{t}}$ and $X_{\tilde{j}\delta}$. We can express τ, (defined in (2.2)), as

$$
\tau \;=\; -\log\left(\frac{1}{\delta} \int_{A_{\tilde{j}}} e^{X_u - X_{\tilde{t}}} du\right).
$$

Hence, translational invariance of the law of X implies that τ conditioned on $X_{(\tilde{\jmath}-1)\delta}$, $X_{\tilde{t}}$ and $X_{\tilde{\jmath}\delta}$ is a function only of $X_{\tilde{t}} - X_{(\tilde{\jmath}-1)\delta}$ and $X_{\tilde{t}} - X_{\tilde{\jmath}\delta}$. Therefore

$$
\begin{aligned}
& E\left[\exp\left(\alpha X_1 + \beta X_{\tilde{t}} + \gamma \frac{\tau}{\sqrt{\delta}}\right)\right] \\
= \; & E\left[\exp(\alpha X_1 + \beta X_{\tilde{t}}) h\left(\gamma; \frac{X_{\tilde{t}} - X_{(\tilde{\jmath}-1)\delta}}{\sqrt{\delta}}, \frac{X_{\tilde{t}} - X_{\tilde{\jmath}\delta}}{\sqrt{\delta}}\right)\right] \\
\approx \; & E\left[\exp(\alpha Y_1 + \beta Y_{\tilde{t}}) h\left(\gamma; \frac{X_{\tilde{t}} - X_{(\tilde{\jmath}-1)\delta}}{\sqrt{\delta}}, \frac{X_{\tilde{t}} - X_{\tilde{\jmath}\delta}}{\sqrt{\delta}}\right)\right]
\end{aligned} \tag{B.1}
$$

for some function h, and where Y is defined by

$$
Y_u = \begin{cases} X_u & 0 \le u \le (\tilde{\jmath}-1)\,\delta \\ X_{(\tilde{\jmath}-1)\delta} + B_{u-(\tilde{\jmath}-1)\delta} & (\tilde{\jmath}-1)\,\delta \le u \le \tilde{\jmath}\delta \\ X_{(\tilde{\jmath}-1)\delta} + B_\delta + X_u - X_{\tilde{\jmath}\delta} & \tilde{\jmath}\delta \le u \le 1 \end{cases}
$$

B is a Brownian motion with the same drift and volatility as X but independent of X. So Y is obtained from X by replacing the path over the interval $[(\tilde{\jmath}-1)\,\delta, \tilde{\jmath}\delta]$ with an independent Brownian motion. The expectation in (B.1) splits into a product:

$$
\begin{aligned}
& E\left[\exp(\alpha Y_1 + \beta Y_{\tilde{t}}) h\left(\gamma; \frac{X_{\tilde{t}} - X_{(\tilde{\jmath}-1)\delta}}{\sqrt{\delta}}, \frac{X_{\tilde{t}} - X_{\tilde{\jmath}\delta}}{\sqrt{\delta}}\right)\right] \\
= \; & E\left[\exp(\alpha Y_1 + \beta Y_{\tilde{t}})\right] E\left[h\left(\gamma; \frac{X_{\tilde{t}} - X_{(\tilde{\jmath}-1)\delta}}{\sqrt{\delta}}, \frac{X_{\tilde{t}} - X_{\tilde{\jmath}\delta}}{\sqrt{\delta}}\right)\right]
\end{aligned}
$$

We would expect the difference between the paths X and Y to be small. The difference between $E\left[\exp(\alpha X_1 + \beta X_{\tilde{t}})\right]$ and $E\left[\exp(\alpha Y_1 + \beta Y_{\tilde{t}})\right]$ will therefore also be small and thus the correlation between $g(X, b)$ and τ is small.

Appendix C

Discretised convertible bonds

In this appendix we look at a discrete version of the convertible bonds problem of Chapter 5. In Chapter 5 we assumed (at least when m exceeded ε) that there was a continuum of infinitesimal bondholders who could not collaborate. We now change the model. We suppose that there are a finite number of bondholders each holding ε units of bonds, and that bonds can only be converted in units of ε. We retain the assumption that the bonds are converted sequentially — if several bondholders all wish to convert, one is chosen at random and he converts his ε bonds. We also retain the assumption that the bondholders do not collaborate. It is clear that the process m_t now takes values in the set $\mathcal{U}$ defined by

$$\mathcal{U} = \{j\varepsilon \; : \; j \in \mathbb{N}, \; j\varepsilon < n\}.$$

The structure of the rest of this appendix is as follows. Conjecture C.1 is analogous to Theorem 5.1 and claims the existence of a solution to the problem. The functions S and B which will correspond to the share and bond price functions are defined in terms of the boundaries $\xi(m)$ and $\eta(m)$ in definition C.1. There are only minor differences from definition 5.1. Finally, Theorem C.1 gives sufficient properties of the functions S and B that they form a Nash equilibrium. Theorem C.1 is analogous to Theorem 5.2.

Conjecture C.1. *There exist functions $\xi : \mathcal{U} \to \mathbb{R}^+$ and $\eta : \mathcal{U} \to \mathbb{R}^+$ satisfying the* *inequalities*

$$\begin{aligned}\xi(m) &< \frac{m\rho'}{\delta} &\quad (C.1)\\ \xi(m) &< \eta(m) &\quad (C.2)\\ \frac{n\rho}{\delta} &\leq \eta(\varepsilon) &\quad (C.3)\end{aligned}$$

such that

(i) η *is decreasing*

in terms of which the optimal policies are

Π_S: *the management of the firm defaults when* $V_t \leq \xi(m_t)$

Π_B: *the bondholders convert when* $V_t \geq \eta(m_t)$.

We now define the functions S and B in terms of the default and conversion boundaries.

Definition C.1. *Given boundaries* ξ *and* η *which satisfy the inequalities (C.1)-(C.3) and condition (i) of conjecture C.1, construct* $\tilde{S} : \mathcal{U} \times \mathbb{R}^+ \to \mathbb{R}$ *by*

(S1) $\tilde{S}(m, V) = 0$ *for* $V \leq \xi(m)$

(S2) $\mathcal{L}\tilde{S} + \dfrac{\delta V - m\rho'}{n - m} = 0$ *for* $V \in [\xi(m), \eta(m)]$

(S3) $V \mapsto \tilde{S}(m, V)$ *is* C^1 *at* $\xi(m)$

and $\tilde{B} : \mathcal{U} \times \mathbb{R}^+ \to \mathbb{R}$ *by*

(B1) $\tilde{B}(m, V) = pV/m$ *for* $V \leq \xi(m)$

(B2) $\mathcal{L}\tilde{B} + \rho = 0$ *for* $V \in [\xi(m), \eta(m)]$

(B3) $\tilde{B}(m, \eta(m)) = \tilde{S}(m, \eta(m))$

(B4) $V \mapsto \tilde{B}(m, V)$ *is continuous at* $\xi(m)$ *and* $\eta(m)$.

Now define $\zeta : [0, \eta(\varepsilon)] \to [\varepsilon, n]$ *by*

$$\zeta(V) \quad = \quad \sup\{\varepsilon \le m < n : \ \eta(m) \ge V\}$$

and S *and* B *by*

$$\begin{aligned}
S(m,V) &= \tilde{S}(m \wedge \zeta(V), V) && \text{(C.4)}\\
B(m,V) &= \tilde{B}(m,V) \quad \textit{for } 0 \le V \le \eta(m) && \text{(C.5)}\\
B(m,V) &= \frac{\varepsilon}{m} S(m-\varepsilon, V) + \frac{m-\varepsilon}{m} B(m-\varepsilon, V) \quad \textit{for } \eta(m) < V \le \eta(\varepsilon). && \text{(C.6)}
\end{aligned}$$

The function $S(0,V)$ gives the share price when no bonds remain, so $S(0,V) = V/n$. As in chapter 5, we interpret $B(0,V)$ also as the share price when no bonds remain. This definition of S and B is very similar to that in definition 5.1. The remarks that follow definition 5.1 apply here too. The following theorem is analogous to Theorem 5.2 and gives sufficient conditions for a Nash equilibrium.

Theorem C.1. *Suppose we are given boundaries* $\xi(m)$ *and* $\eta(m)$ *which satisfy the inequalities (C.1)-(C.3) and condition (i) of conjecture C.1. If the functions* S *and* B *of definition C.1 have the property*

$(P1)'$ $S(m, \eta(m)) = S(m-\varepsilon, \eta(m)) = B(m-\varepsilon, \eta(m))$

then we extend S *and* B *continuously to* $\mathcal{U} \times \mathbb{R}^+$ *by*

$(P2)'$ $B(m,V) = S(m,V) = V/n$ *for* $V \ge \eta(\varepsilon)$.

Suppose that the functions S *and* B *so extended have the properties*

$(P3)'$ $V \mapsto B(\varepsilon, V)$ *is* C^1 *at* $\eta(\varepsilon)$

$(P4)'$ $B(\varepsilon, V) \ge V/n$

$(P5)'$ $S(m-\varepsilon, V) < B(m,V)$ *for* $V < \eta(m)$

$(P6)'$ $B(m-\varepsilon, V) \le B(m,V)$ *for* $V \ge \eta(m)$

$(P7)'$ $S(m,V) \ge 0$

then S corresponds to the share price, B corresponds to the bond price, and the policies Π_S and Π_B are optimal.

Proof of Theorem C.1. This proof is analogous to that of Theorem 5.2. Define the local martingale X_t by

$$\begin{aligned} X_t &= e^{-rt}S(m_t, V_t) - \int_0^t e^{-ru}\mathcal{L}S(m_u, V_u)du \\ &\quad - \sum_{0<u\leq t} e^{-ru}\{S(m_u, V_u) - S(m_{u-}, V_{u-})\}. \end{aligned}$$

For any stopping time T that reduces X we have

$$\begin{aligned} X_0 &= S(m_0, V_0) = E[X_T] \\ &= E\Bigg[e^{-rT}S(m_T, V_T) - \int_0^T e^{-ru}\mathcal{L}S(m_u, V_u)du \\ &\quad - \sum_{0<u\leq T} e^{-ru}\{S(m_u, V_u) - S(m_{u-}, V_{u-})\}\Bigg]. \end{aligned}$$

Now under Π_B conversion occurs only when $V \geq \eta(m)$. Hence from (C.4), *(P1)′* and *(P2)′* we see that S does not change at conversion. Therefore

$$S(m_0, V_0) = E\left[e^{-rT}S(m_T, V_T) - \int_0^T e^{-ru}\mathcal{L}S(m_u, V_u)du\right].$$

The same argument that was used in the proof of Theorem 5.2 shows that

(i) $S(m_0, V_0)$ is an upper bound for the net present value of the future cashflows from a share,

(ii) Under Π_S, $S(m_0, V_0)$ equals the net present value of the future cashflows from a share.

Therefore $S(m, V)$ is the share price. We do not repeat the details of this argument.

Now let $\bar{T}$ denote the time at which the management of the firm defaults. We assume that the management of the firm follows Π_S, all bondholders follow Π_B, and $V_0 \leq$

$\eta(m_0)$. We consider the last bondholder to convert. Define the local martingale Z_t by

$$\begin{aligned} Z_t &= e^{-r(t\wedge\bar{T})}B(m_{t\wedge\bar{T}}, V_{t\wedge\bar{T}}) - \int_0^{t\wedge\bar{T}} e^{-ru}\mathcal{L}B(m_u, V_u)du \\ &- \sum_{0<u\leq t\wedge\bar{T}} e^{-ru}\left\{B(m_u, V_u) - B(m_{u-}, V_{u-})\right\}. \end{aligned}$$

If T reduces Z we obtain

$$\begin{aligned} Z_0 &= B(m_0, V_0) = E[Z_T] \\ &= E\left[e^{-r(T\wedge\bar{T})}B(m_{T\wedge\bar{T}}, V_{T\wedge\bar{T}}) - \int_0^{T\wedge\bar{T}} e^{-ru}\mathcal{L}B(m_u, V_u)du\right. \\ &- \left.\sum_{0<u\leq T\wedge\bar{T}} e^{-ru}\left\{B(m_u, V_u) - B(m_{u-}, V_{u-})\right\}\right]. \end{aligned}$$

Conversion will occur when $V_t = \eta(m_t)$ and thus *(B3)'* and *(P1)'* show that B does not change at conversion. Therefore

$$B(m_0, V_0) = E\left[e^{-r(T\wedge\bar{T})}B(m_{T\wedge\bar{T}}, V_{T\wedge\bar{T}}) - \int_0^{T\wedge\bar{T}} e^{-ru}\mathcal{L}B(m_u, V_u)du\right].$$

Again, it follows exactly as in the proof of Theorem 5.2 that $B(m_0, V_0)$ is the bond price.

Next consider a bondholder who converts before $m = \varepsilon$ under Π_B. Properties *(B3)'* and *(P1)'* show that the net present value of his future cashflows when he converts does not change. Thus $B(m_0, V_0)$ also defines the net time-0 value of his cashflow.

If $V_0 > \eta(m_0)$ then under Π_B bonds are converted immediately. Each bondholder will attempt to convert; with probability ε/m a specific bondholder will be the one chosen to convert, with probability $(m-\varepsilon)/m$ he will remain a bondholder. Thus (C.6) shows that $B(m, V)$ corresponds to the bond price.

We now show the suboptimality of alternative policies. If a bondholder chooses to convert when $V_t < \eta(m_t)$ then he is swapping from the cashflow associated with a bond to that associated with a share. Property *(P5)'* shows that the net present value of his future cashflows decreases. Thus his behaviour is suboptimal.

The other way in which a bondholder could change policy is by not attempting to convert when $V \geq \eta(m)$. If $m = \varepsilon$ then the proof that this strategy is suboptimal follows as in the proof of Theorem 5.2. If $m > \varepsilon$ then the other bondholders will attempt to convert. It follows from *(P6)'* that this alternative policy for the bondholder is suboptimal. □

We have conjectured that a solution exists. We now outline a method for finding this solution. First observe that $\xi(\varepsilon)$ and $\eta(\varepsilon)$ can be constructed as described in lemma 5.2. It is therefore sufficient to describe how to obtain the boundaries $\xi(m)$ and $\eta(m)$ from $\xi(m-\varepsilon)$ and $\eta(m-\varepsilon)$. Recall the definition of Y:

$$Y(m, V) \quad = \quad S(m, V) - B(m, V).$$

From property *(P1)'* we deduce that $\eta(m)$ is a root of

$$Y(m-\varepsilon, V) \quad = \quad 0. \tag{C.7}$$

Condition (i) of conjecture C.1 states that $\eta(m) \leq \eta(m-\varepsilon)$. Furthermore, property *(B3)'* implies that $\eta(m-\varepsilon)$ is a root of (C.7). As $Y(m-\varepsilon, \xi(m-\varepsilon))$ is negative it follows from lemma 5.1 that there is at most one root to (C.7) on the interval $(\xi(m-\varepsilon), \eta(m-\varepsilon))$. We define $\eta(m)$ to be this root if it exists, otherwise we choose $\eta(m) = \eta(m-\varepsilon)$. The value of $\xi(m)$ is then chosen so that $V \mapsto S(m, V)$ is continuous at $\eta(m)$. It is possible to prove *(P7)'* in a similar way that the corresponding condition was proved in section 5.3. It is also possible to prove *(P6)'*. However, conjecture C.1 remains a conjecture owing to the difficulty in proving *(P5)'*. This, and the difficulties in verifying *(P5)'* numerically, are discussed in section 5.7.

Bibliography

Arrow, K.J., & Kurz, M. 1970. *Public Investment, the Rate of Return, and Optimal Fiscal Policy.* John Hopkins University Press, Baltimore.

Bick, A. 1987. On the Consistency of the Black-Scholes Model with a General Equilibrium Framework. *Journal of Financial and Quantitative Analysis*, **22**, 259–275.

Black, F., & Scholes, M. 1973. The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, **81**, 637–654.

Border, K.C. 1985. *Fixed Point Theorems with Applications to Economics and Game Theory.* Cambridge University Press.

Breeden, D.T. 1986. Consumption, Production, Inflation and Interest Rates: a Synthesis. *Journal of Financial Economics*, **16**, 3–39.

Brennan, M.J., & Schwartz, E. 1977. Convertible Bonds: Valuation and Optimal Strategies for Call and Conversion. *Journal of Finance*, **32**, 1699–1715.

Brennan, M.J., & Schwartz, E. 1980. Analyzing Convertible Bonds. *Journal of Financial and Quantative Analysis*, **15**, 907–929.

Brennan, M.J., & Schwartz, E. 1989. Portfolio Insurance and Financial Market Equilibrium. *Journal of Business*, **62**, 455–476.

Broadie, M., Glasserman, P., & Kou, S. 1997. A Continuiuty Correction for Discrete Barrier Options. *Mathematical Finance*, **4**, 325–348.

Carr, P. 1995. Two Extensions to Barrier Option Valuation. *Applied Mathematical Finance*, **2**, 173–209.

Constantinides, G.M. 1984. Warrant Exercise and Bond Conversion in Competitive Markets. *Journal of Financial Economics*, **13**, 371–397.

Constantinides, G.M., & Rosenthal, R.W. 1984. Strategic Analysis of the Competitive Exercise of Certain Financial Options. *Journal of Economic Theory*, **32**, 128–138.

Conze, A., & Viswanathan. 1991. Path Dependent Options: The Case of Lookback Options. *Journal of Finance*, **46**, 1893–1907.

Diamond, P.A. 1984. Money in a Search Equilibrium. *Econometrica*, **52**, 1–20.

Duffie, D., Gârleanu, N., & Pedersen, L.H. 2000. *Valuation in Dynamic Bargaining Markets.* Graduate School of Business, Stanford University, working paper.

Emanuel, D.C. 1983. Warrant Valuation and Exercise Strategy. *Journal of Financial Economics*, **12**, 211–235.

Ethier, S.N., & Kurtz, T.G. 1986. *Markov Processes: Characterization and Convergence.* Wiley.

Frey, R., & Stremme, A. 1997. Market Volatility and Feedback Effects from Dynamic Hedging. *Mathematical Finance*, **7**, 351–374.

Gennotte, G., & Leland, L. 1990. Market Liquidity, Hedging and Crashes. *The American Economic Review*, **80**, 999–1021.

Guckenheimer, J., & Holmes, P. 1983. *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields.* Springer-Verlag.

Harrison, J.M., & Kreps, D.M. 1979. Martingales and Arbitrage in Multiperiod Security Markets. *Journal of Economic Theory*, **20**, 381–408.

Harrison, J.M., & Pliska, S.R. 1981. Martingales and Stochastic Integrals in the Theory of Continuous Trading. *Stochastic Processes and their Applications*, **11**, 215–260.

Heritage, J.P. 2002. Pricing Moving Average Barrier Options. *Journal of Computational Finance*, **5**, 51–67.

Heritage, J.P., & Rogers, L.C.G. 2002. Large Investors, Takeovers, and the Rule of Law. *Monte Carlo Methods and Applications*, **8**, 357–370.

Ingersoll, J.E. 1977. A Contingent-Claims Valuation of Convertible Securities. *Journal of Financial Economics*, **4**, 289–321.

Itô, K., & McKean, H.P. 1996. *Diffusion Processes and their Sample Paths.* Springer-Verlag.

Karatzas, I., Lehoczky, J.P., & Shreve, S.E. 1987. Optimal Portfolio and Consumption Decisions for a "Small Investor" on a Finite Horizon. *SIAM Journal on Control and Optimization*, **25**, 1557–1586.

Karatzas, I., Lehoczky, J.P., & Shreve, S.E. 1990. Existence and Uniqueness of Multi-Agent Equilibrium in a Stochastic, Dynamic Consumption/Investment Model. *Mathematics of Operations Research*, **15**, 80–128.

Kiyotaki, N., & Wright, R. 1991. A Contribution to the Pure Theory of Money. *Journal of Economic Theory*, **53**, 215–235.

Kiyotaki, N., & Wright, R. 1993. A Search-Theoretic Approach to Monetary Economics. *American Economic Review*, **83**, 63–77.

Kydland, F.E., & Prescott, E.C. 1977. Rules Rather than Discretion: The Inconsistency of Optimal Plans. *Journal of Political Economy*, **85**, 473–491.

Leland, H.E. 1994. Corporate Debt Value, Bond Covenants, and Optimal Capital Structure. *Journal of Finance*, **49**, 1213–1252.

Leland, H.E., & Toft, K.B. 1996. Optimal Capital Structure, Endogenous Bankruptcy, and the Term Structure of Interest Rates. *Journal of Finance*, **51**, 987–1019.

Merton, R.C. 1969. Lifetime Portfolio Selection Under Uncertainty: The Continuous-Time Case. *Review of Economics and Statistics*, **51**, 247–257.

Merton, R.C. 1971. Optimal Consumption and Portfolio Rules in a Continuous-Time Model. *Journal of Economic Theory*, **3**, 373–413.

Merton, R.C. 1973. Theory of Rational Option Pricing. *Bell Journal of Economics and Management Science*, **4**, 141–183.

Merton, R.C. 1974. On the Pricing of Corporate Debt: The Risk Structure of Interest Rates. *Journal of Finance*, **29**, 449–470.

Musiela, M., & Rutkowski, M. 1998. *Martingale Methods in Financial Modelling*. Springer.

Nyborg, K.G. 1996. The Use and Pricing of Convertible Bonds. *Applied Mathematical Finance*, **3**, 167–190.

Platen, E., & Schweizer, M. 1998. On Feedback Effects from Hedging Derivatives. *Mathematical Finance*, **8**, 67–84.

Revuz, D., & Yor, M. 1991. *Continuous Martingales and Brownian Motion.* Springer-Verlag.

Rogers, L.C.G., & Satchell, S.E. 1991. Estimating Variance from High, Low and Closing Prices. *Annals of Applied Probability*, **1**, 504–523.

Rogers, L.C.G., & Williams, D. 2000. *Diffusions, Markov Processes and Martingales.* Vol. 2. Cambridge University Press.

Rogers, L.C.G., & Yousaf, F.A. 2000. *Changes in Market Equilibria Caused by Sudden Shocks.* University of Bath, preprint.

Rubinstein, M., & Reiner, E. 1991. Breaking Down the Barriers. *RISK*, **4**(September), 28–35.

Samuelson, P.A. 1969. Lifetime Portfolio Selection by Dynamic Stochastic Programming. *Review of Economics and Statistics*, **51**, 239–246.

Shi, S. 1995. Money and Prices: A Model of Search and Bargaining. *Journal of Economic Theory*, **67**, 467–496.

Trejos, A., & Wright, R. 1995. Search, Bargaining, Money, and Prices. *Journal of Political Economy*, **103**, 118–141.

Vervaat, W. 1979. A Relation between Brownian Bridge and Brownian Excursion. *Annals of Probability*, **7**, 143–149.